(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
24 June 2004 (24.06.2004)

PCT

(10) International Publication Number
WO 2004/053137 A2

(51) International Patent Classification[7]: C12N 15/85

(21) International Application Number:
PCT/CH2003/000810

(22) International Filing Date:
10 December 2003 (10.12.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/432,245 11 December 2002 (11.12.2002) US

(71) Applicant *(for all designated States except US)*: **CYTOS BIOTECHNOLOGY AG** [CH/CH]; Wagistrasse 25, CH-8952 Schlieren (CH).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **SAUDAN, Philippe** [CH/CH]; Dürrenrainstrasse 20, CH-8422 Pfungen (CH). **HENNECKE, Frank** [DE/CH]; Aufwiesenstrasse 14, CH-8305 Dietlikon (CH). **BACHMANN, Martin, F.** [CH/CH]; Goldackerweg 8, CH-8472 Seuzach (CH). **STERN, Dorothee, M.** [DE/CH]; In der Fadmatt 56, CH-8902 Urdorf (CH).

(74) Common Representative: **CYTOS BIOTECHNOLOGY AG**; c/o Sperrle, Martin, Wagistrasse 25, CH-8952 Schlieren (CH).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 2004/053137 A2

(54) Title: METHOD FOR PROTEIN PRODUCTION

(57) Abstract: The present invention relates to the field of gene expression technology and the production of recombinant polypeptides and/or untranslated RNA molecules in host cells. In particular, the invention provides a method for the production of recombinant polypeptides and/or untranslated RNA molecules in host cells allowing the production of high amounts of those molecules in a very short time and with a high degree of reproducibility, in particular with respect to a, preferably, pre-defined standard and, thus, making the present invention particularly attractive for the production of proteins for clinical use and the production under Good Manufacturing Practice (GMP).

METHOD FOR PROTEIN PRODUCTION

BACKGROUND OF THE INVENTION

The present invention relates to the field of gene expression technology and the production of recombinant polypeptides and/or untranslated RNA molecules in host cells. In particular, the invention provides a method for the production of recombinant polypeptides and/or untranslated RNA molecules in host cells allowing the production of high amounts of those molecules in a very short time and with a high degree of reproducibility, preferably with respect to a pre-defined standard and, thus, making the present invention particularly attractive for the production of proteins for clinical use and the production under Good Manufacturing Practice (GMP).

Related Art

Process development for biopharmaceuticals, and hereby in particular the expression of polypetides used in these biopharmaceuticals, is governed by the economy of the manufacturing process. The classical approach for the production of recombinant polypetides is the use of expression systems based on chromosomal integration of an expression plasmid into the genome of the host cell. The sites of gene integration, however, are random, and the number and ratio of genes integrating at any particular site are unpredictable. As a result, it is normally necessary to screen many clonal cell populations to obtain a cell line in which the desired genes are expressed at an appropriate level. This procedure of transfection, selection and analysis of numerous clonally derived cell lines expressing the multiple genes is costly and time-consuming.

An alternative to expression systems based on chromosomal integration of an expression plasmid into the genome of the host cell are transient expression systems. The expression of the latter systems is based on non-integrated plasmids, and hence the expression is typically rapidly lost as the cell undergoes division. Thus, transient expression systems typically do not lead to sufficient expression over time implying that repeated processes would be necessary which might not be desirable.

Transient expression can be performed by infection (using, e.g., viral vectors like adenoviral vectors) or by transfection of the appropriate vector using techniques well known in the art, such as e.g polyethyleneimine (PEI), calcium phosphate coprecipitation method or continous flow electroporation. Transfection can be performed with either a non-replicating or a replicating DNA vector or with a virus derived RNA vector (Blasey et al., 1999, in: Animal cell technology, 331-337). A variety of virus vectors for expression in higher eukaryotic cells have been developed by using recombinant DNA technology and introducing genes of interest into virus genomes. The initial viruses used in vector development were DNA viruses such as Adeno-associated viruses or Poxviruses as well as RNA viruses replicating through DNA intermediates such as Retroviruses.

Another way of performing transient expression is by transfection of an episomal vector which replicates extra-chromosomally in a host cell such as in a mammalian host cell. Episomal replicating vectors provide certain advantages over classical expression systems since they do not integrate into the host genome, but replicate episomally in a host cell. Most of the described episomally replicating vectors are based on viral components from viruses like Papovaviridae (e.g. SV40, BPV) or Herpesviridae (e.g. EBV). Episomal replicating vectors derived from these viruses generally contain a replication origin and at least one viral trans-acting factor, e.g., an initiator protein. Examples of such initiator proteins are large T-antigen for SV40, E1/E2 for BPV, and EBNA-1 for EBV. The process of episomal replication typically involves both host cell replication machinery and virus trans-acting factors (Bode et al., 2001 Gene Ther Mol Biol, v6, 33-46).

Transient gene expression in, e.g., mammalian cells, at reactor scale is becoming increasingly important for the rapid production of recombinant proteins. Large scale applications of transient expression systems for recombinant protein production have been reported (Wurm and Bernard, 1999, Curr Opin Biotechn, 10, 156-159; Meissner et al., 2001, Biotechn and Bioeng, 75, 2, 197-203; Durocher et al., 2002, Nucleic Acid Res, 30, 1-9).

Large-scale transient gene expression allows expression of recombinant proteins, typically, in less than one month making such a process faster as compared to expression systems based on chromosomal integration of an expression plasmid into the genome of the host cell. However, common to these processes is the generation of rather large cell populations prior to transfection requiring a huge amount of DNA which has to be

prepared for large scale transient transfections. Moreover, a complete medium exchange prior to transfection is often required as the presence of cellular by-products in conditioned medium is associated with poor transfection yield (Belting et al., 1999, J. Biol. Chem, 274, 19375-19382). The requirement of the huge amount of DNA which has to be prepared for large scale transient transfection is one major disadvantage of this process. Furtheron transfection of serum-free adapted cell lines is often less efficient then in adherent cell lines (Jordan et al., 1998, Cytotechnology, 26, 39-47).

One other disadvantage of large-scale transient expression systems known in the art lies in the transient nature of the process itself. Once the production phase is finished a new process has to be started right from the beginning. Variations in the transfection efficiency of large scale processes typically leads to a high variation in the rate of yield of the recombinant proteins and, thus, leads to a low degree of reproducibility of those processes which is, in particular, unfavourable for the production of proteins for clinical use and the production under GMP (Good Manufacturing Practice).

As with the tools of genomics and proteomics a vast number of proteins of therapeutic interest are identified, the requirement for efficient and rapid expression of genes is increasing. Cost-effective manufacturing processes that allow rapid expression of the desired protein/product in high amounts for characterization, high throughput production and production for clinical applications under Good Manufacturing Practice (GMP) conditions are an essential tool.

In particular, for the purpose of producing recombinant polypeptides and/or untranslated RNA molecules for clinical use and for the production under GMP, a high reliability and reproducibility of those processes are required. The current available processes for protein production do not provide sufficient reliability and reproducibility to allow setting up a complete predeveloped expression process independent of the polypeptide or untranslated RNA which has to be expressed. This is in particular of major interest for a process for protein production under Good Manufacturing Practice (GMP) on a commercial and large scale.

Therefore, a need exists in the art to develop a well characterized process for the production of recombinant proteins under GMP conditions with a high degree of reproducibility, preferably with respect to a pre-defined standard that allows rapid expression of any desired product in high amounts.

Moreover, there is a need for a method that avoids large-scale transient transfection with low transfection efficiencies and that avoids the use of large amounts of DNA and transfection agents. Therefore, a method is needed that allows transfection of cells in small scale as adherent cultures to assure high transfection efficiencies, and a method that requires low amounts of DNA and transfection agent and that provides high reproducibility and reliability which furthermore allows the production of proteins for clinical use and the production under GMP.

SUMMARY OF THE INVENTION

The present invention provides methods that allow the production of polypeptides and/or untranslated RNA molecules. More specifically, the invention provides methods for the production of polypeptides and/or untranslated RNA molecules which are, in particular, suitable for use under GMP conditions with a high degree of reproducibility, preferably with respect to a pre-defined standard. The methods of the invention are also useful, e.g., for the isolation and/or purification of polypeptides and/or untranslated RNA molecules that are produced in host cells.

In certain embodiments, the invention provides methods for regulated, in particular temperature inducible, expression of polypeptides and/or untranslated RNA molecules in recombinant host cells, preferably by way of transfection, making the methods of the present invention suitable for the production of cytotoxic polypeptides, e.g. polypeptides that are detrimental to the viability of a host cell. The present invention also provides methods which allow rapid and high level production of specific RNA molecules produced in transfected, recombinant host cells.

In a general aspect, the present invention provides for a method for producing a polypeptide or untranslated RNA molecule, the method comprising the steps of (A) providing at least one host cell; (B) introducing at least one nucleic acid molecule into the at least one host cell to produce at least one recombinant host cell, the one or more nucleic acid molecule comprising (a) a first polynucleotide element capable of replicating the at least one nucleic acid molecule in the at least one host cell; and (b) at least one second polynucleotide element selected from the group consisting of (i) an open reading frame encoding a polypeptide of interest; (ii) a nucleotide sequence complementary to all or a part of the open reading frame of (i); and (iii) a nucleotide sequence encoding an

untranslated RNA molecule or complement thereof; (C) selecting at least one stable recombinant host cell from the at least one host cell, (D) culturing the recombinant host cell under conditions suitable for expression of the polypeptide or untranslated RNA molecule; and (E) controlling the reproducibility of the method for producing the polypeptide or the untranslated RNA molecule as a function of time as compared to a pre-defined standard.

In certain embodiments, controlling the reproducibility of the method of the invention compared to a pre-defined standard is effected by way of measuring the amount of the polypeptide or untranslated RNA molecule per cell expressed by the at least one recombinant host cell as a function of time. Alternatively, the controlling the reproducibility of the method is effected by way of determining the number of the nucleic acid molecule being in the at least one recombinant host cell as a function of time. The determination is, preferably, performed by quantitative PCR (polynucleotide chain reaction) or alternatively, by southern blot.

In a further preferred aspect of the present invention, the host cell of the present invention is adapted to serum-free growth. In another aspect, the host cell is adapted to serum-free growth prior to the introducing the nucleic acid molecule into the host cell to produce a recombinant host cell, that is prior to transfection with the nucleic acid molecule of the invention. In one embodiment of the invention, the host cell is a eukaryotic host cell. In an alternative embodiment, the host cell is a prokaryotic host cell.

The at least one host cell may be one or more host cell, preferably a host cell line, in particular a culture of a particular type of cell that can be reproduced, typically, indefinitely. In another embodiment, more than one host cell line may be provided for the method of the present invention.

The methods of the invention may further comprise recovering the polypeptide or untranslated RNA molecule.

The methods of the present invention, thus, allow for transfection, selection, and generation of stable recombinant host cell populations, and for very rapid large-scale production of polypeptides. The invention provides well characterized processes with high reliability and reproducibility, preferably with respect to a pre-defined standard, that allow setting up a complete predeveloped expression process independent of the polypeptide or untranslated RNA which has to be expressed, therefore making the methods useful for producing recombinant polypeptides and/or untranslated RNA

molecules for clinical use and for protein production under Good Manufacturing Practice (GMP) on a commercial scale and that allows rapid expression of the desired product in high amounts. One embodiment and hereby the inducibility of the vector makes the invention especially suitable for the production of cytotoxic polypeptides, e.g., polypeptides that are detrimental to the viability of a host cell.

Therefore, the methods of the present invention allow, typically and preferably, transfection of cells in small scale as adherent cultures to assure high transfection efficiencies, and that require low amounts of DNA and transfection. Thus, preferably, stable recombinant host cell populations are generated in small scale that are scaled up after selection of the at least one stable recombinant host cell for large-scale production of polypeptides. In particular, the method of the invention for the production of polypeptides or untranslated RNA is a very economical process that allows rapid expression of the desired protein/product in high amounts for characterization and high throughput production, and that is therefore suitable for the production for clinical applications and for use under Good Manufacturing Practice (GMP) conditions with a high degree of reproducibility, preferably with respect to a pre-defined standard, that is to say for protein production conforming to GMP.

When the terms "one," "a," or "an" are used in this disclosure, they mean "at least one" or "one or more," unless otherwise indicated.

BRIEF DESCRIPTION OF THE FIGURES

Fig. 1 is a schematic representation of vector pCEPpu described in Example 1. The hygromycin resistance marker in pCEP4 was exchanged by the puromycin resistance marker. The EGFP and EPO expression cassette was cloned into the multiple cloning site (MCS) of vector pCEPpu. MCS: multiple cloning site; CMV: Cytomegalovirus promoter, pA: SV40 polyadenylation signal; puro: puromycin resistance marker; oriP: Epstein-barr virus origin of replication; EBNA1: replication initiation factor EBNA-1; amp: ampicillin resistance marker; ColE1: bacterial origin of replication.

Fig. 2 shows EGFP expression in different cell lines transfected with vector pCEP-EGFP. A stable cell pool was selected in the presence of puromycin and kept at 37°C over

several weeks. EGFP expression was determined by flow cytometry at different time points.

Fig. 3 shows EPO expression in 293EBNA cells transfected with vector pCytTS-EPO-OriP. Selected cell populations were either kept at 37°C or shifted to 29°C in order to induce the Sindbis replicon. Cell culture supernatant was collected from uninduced and induced cells seven days after induction. EPO levels were quantified by ELISA.

Fig. 4 shows the rapid expression of EPO in 293EBNA cells in batch fermentation. Serum-free adapted 293EBNA cells were transfected with pCytTs-EPO-OriP. After selection the stable cell pool was expanded in serum-free medium to a total volume of 8900 ml in a 10 l Wave® reactor bag. Cells were shifted to 29°C to induce EPO expression. EPO levels were quantified by ELISA. The induction phase is shown by a grey box.

Fig. 5 shows the rapid expression of EPO in 293EBNA cells in perfusion fermentation. Serum-free adapted 293EBNA cells were transfected with pCytTs-EPO-OPE. After selection the stable cell pool was expanded in serum-free medium to a total volume of 1000 ml in a 1 l Wave® reactor bag. A perfusion process was performed over six days with medium exchange twice a day. The reactor was induced at 29°C and EPO expression was monitored over 4 days after induction. The perfusion phase of the process is shown by a hatched box and the induction phase by a grey box.

DETAILED DESCRIPTION OF THE INVENTION

The method of the invention and the nucleic acid molecules, expression systems and recombinant host cells of the invention may be used for the production of proteins, polypeptides and RNA molecules, e.g., untranslated RNA molecules. According to one aspect of the invention, a method is provided for producing a polypeptide or untranslated RNA molecule, the method comprising the steps of (A) providing at least one host cell; (B) introducing at least one nucleic acid molecule into the at least one host cell to produce at least one recombinant host cell, the one or more nucleic acid molecule comprising (a) a first polynucleotide element capable of replicating the at least one nucleic acid molecule in the at least one host cell; and (b) at least one second polynucleotide element selected from the group consisting of (i) an open reading frame encoding a polypeptide of interest;

(ii) a nucleotide sequence complementary to all or a part of the open reading frame of (i); and (iii) a nucleotide sequence encoding an untranslated RNA molecule or complement thereof; (C) selecting at least one stable recombinant host cell from the at least one host cell, (D) culturing the recombinant host cell under conditions suitable for expression of the polypeptide or untranslated RNA molecule; and (E) controlling the reproducibility of the method for producing the polypeptide or the untranslated RNA molecule as a function of time, typically and preferably as compared to a pre-defined standard.

In a very preferred embodiment, the steps of the method for producing a polypeptide or untranslated RNA molecule are performed consecutively. Thus, the preferred order of the steps of the method is such that step A is followed by B, then C, D, and E. The introduction of the at least one nucleic acid molecule or vector, respectively, into the at least one host cell (step B) prior to selection (step C) and cultivation (step D) of said at least one recombinant host cell assures high transfection efficiencies, and requires low amounts of the at least one nucleic acid molecule or vector, respectively. Thus, preferably, the at least one stable recombinant host cell is cultured (step D) after selection of the at least one stable recombinant host cell under conditions suitable for expression and for large-scale production of the polypeptide or untranslated RNA molecule.

The term "small scale" that is used with reference to the step of introducing at least one nucleic acid molecule into at least one host cell (step B), preferably with reference to transfection of the at least one nucleic acid molecule into at least one host cell, is intended to mean that the at least one host cell provided for the first step (step A) of the method of the invention is preferably a small cell culture, preferably a culture that assures high transfection efficiencies, preferably a culture with the total cell number which is ten times less then the total cell number after the cultivation step (step D).

The term "large-scale" with reference to the expression or production of a polypeptide or untranslated RNA molecule is intended to mean that the final total cell number of the host cell population after cultivation (step D) is at least ten times higher than the total cell number that is provided for the first step (step A) of the method of the invention. The total number of cells can be measured by methods well known in the art, such as for example by a cell culture counter.

In one embodiment of the invention the selected at least one stable recombinant host cell is cryopreserved after performing step C. To ensure reproducibility and reliability of the method, the cryopreserved at least one stable recombinant host cell is

rethawed for culturing the recombinant host cell under conditions suitable for expression of the polypeptide or untranslated RNA molecule (step D).

As used herein, the phrase "stable recombinant host cell" refers to a host cell which semi-stably expresses one or more nucleic acid molecules or expression vectors, respectively, of the invention that have been introduced into the host cell, wherein the nucleic acid molecule or expression vector, respectively, comprises at least one selection marker of the invention. Such "stable recombinant host cells" which semi-stably express one or more nucleic acid molecules or expression vectors, respectively, of the invention, continuously produce the protein of interest for longer than 72 hours after introduction, preferably for longer than one week, for longer than one month, for longer than two months, for longer than three months, or even for longer than six months.

The inventive method further comprises selecting stable recombinant host cells which semi-stably express the nucleic acid molecule, preferably the replication-competent expression vector, and maintaining stable recombinant host cells.

The inventive method further comprises culturing the stable recombinant host cells under conditions suitable for expression of the polypeptide or untranslated RNA molecule of the invention. In a further preferred aspect of the present invention, the host cell of the present invention is adapted to serum-free and/or protein free growth, preferably prior to the introduction of the nucleic acid molecule or vector, respectively, into the host cell to produce a recombinant host cell, that is prior to transfection with the nucleic acid molecule of the invention. As used herein, the term 'serum-free' as applied to media includes any cell culture medium that does not contain and being free from, respectively, serum such as e.g. fetal bovine serum and, thus, serum-free growth refers to cell cultures grown in serum-free medium. Similarly, the term 'protein-free' as applied to media includes medium free from exogenously added protein and, thus, protein-free growth refers to cell cultures grown in medium free from exogenously added protein.

In one embodiment of the invention, the host cell is a eukaryotic host cell. In an alternative embodiment, the host cell is a prokaryotic host cell. In certain embodiments, the recombinant host cells may be cultured in reactors that include without limitation bioreactor (e.g. stirred tank, perfused, membrane enclosed, encapsulated cell, fluidized bed, hollow-fiber reactor, airlift reactor), shaker flask, cell factory or roller bottle, porous microbeads, batch reactor, fed-batch reactor, spinner flask, or Wave® reactor. As used herein, the term "bioreactor" means large scale cultivation of cells in specialized reaction

vessels. The term "batch reactor" means cultivation of cells in a bioreactor without addition of nutritions during the cultivation process. The term "fed batch reactor" means cultivation of cells in a bioreactor with the addition of fresh media and/or nutritions during the cultivation process without continuous removal of bioreactor fluid. The term "continous reactor" means cultivation of cells in a bioreactor with continuously addition of fresh media and/or nutritions and continuously removal of bioreactor fluid. The term "Wave® reactor" means large scale cultivation of cells in disposable reaction vessels.

The overall cell culture process employing nucleic acid molecules and expression systems of the invention for the production of polypeptides and/or untranslated RNA molecules can be implemented in a variety of bioreactor configurations (e.g., stirred tank, perfused, membrane enclosed, encapsulated cell, fluidized bed, and air lift reactors) and scales (from laboratory T flasks to thousands of liters), chosen to accommodate the requirements of the host cell line utilized (e.g., anchorage dependency, O2 concentrations), to maximize the production of expression product, and to facilitate subsequent recovery and purification of expression product. In a preferred embodiment, the cell culture process is performed in a Wave® reactor using a batch or perfusion process (Ohashi et al., Genetic Engineering News, 2001, 21, 40; Singh, Cytotechnol., 1999, 30, 149-158; Singh, Genetic Engineering News, 1999, 19, 23).

Further, a number of different bioprocess parameters can be varied in order to alter the glycosylation pattern of polypeptide products produced by the recombinant host cells of the invention. These factors include medium composition, pH, oxygen concentration, lack or presence of agitation, and, for the case of anchorage dependent cells, the surface provided. Thus, the glycosylation pattern of glycoproteins may be altered by choosing the host cell in which these proteins are expressed in and the conditions under which the recombinant host cells are grown.

In a preferred embodiment, the present invention provides methods to ensure and control the reproducibility and reliability of the method for producing a polypeptide of the invention. Preferably, the controlling of the reproducibility of the inventive method is effected by way of measuring the amount of the polypeptide or untranslated RNA molecule per cell expressed by the at least one recombinant host cell as a function of time. The measurement of the amount of polypeptide or untranslated RNA molecule includes without limitation methods like quantitative dot blot, ELISA, SDS-PAGE, Western blot, Northern blot, enzymatic assays, spectrophotometric assays and other methods well

known in the art. To measure the amount of polypeptide or untranslated RNA molecule as a function of time, to ensure the reproducibility and reliability of the method of the invention, aliquots of the medium supernatant and/or aliquots of the recombinant host cells are withdrawn at regular intervals during the production phase and are analyzed by ELISA, SDS-PAGE, Western blot, Northern blot, enzymatic assays, spectrophotometric assays and other methods well known in the art. As a function of time is intended to mean at regular intervals, preferably once a week, more preferably every third day, more preferably every second day, most preferably daily. Thus, in a preferred embodiment, the amount of polypeptide or untranslated RNA molecule per cell expressed by the at least one recombinant host cell is measured at regular intervals, preferably daily.

In another preferred embodiment, the controlling of the reproducibility of the method is effected by way of determining the number of the at least one nucleic acid molecule or vector, respectively, being in the at least one recombinant host cell as a function of time. A wide variety of techniques suitable for determining the nucleic acid molecules and/or vectors of the invention in host cells are well known to those of skill in the art. Such techniques are well described in Sambrook, J. et al., eds., (MOLECULAR CLONING, A LABORATORY MANUAL, 2nd. edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989)) and may include without limitation quantitative PCR or Southern Blot. To determine the number of the nucleic acid molecule as a function of time to ensure the reproducibility and reliability of the method of the invention aliquots of the recombinant host cells are withdrawn daily during the production phase and are analyzed by methods well known in the art, preferably by quantitiative PCR or southern blot. Thus, preferably, the number of the at least one nucleic acid molecule or vector, respectively, being in the at least one recombinant host cell is determined at regular intervals, preferably daily.

As used herein, the term "to control the reproducibility" and "to ensure the reproducibility", respectively, refers to the reproducibility of the inventive method for producing a polypeptide or untranslated RNA molecule and refers hereby, preferably, to the reproducibility of the inventive method with respect to a pre-defined standard. The reproducibility of the inventive method is controlled as a function of time, wherein preferably the deviation of certain parameters as compared to a typically pre-defined standard is determined as a function of time. Typical and preferred parameters used for controlling the reproducibility are the amount of the polypeptide or untranslated RNA

molecule per cell being expressed by the recombinant host cells or the number of nucleic acid molecules being in the recombinant host cells after its introduction herein. To ensure a minimal deviation with respect to a pre-defined standard and, thus, to ensure a certain degree of reproducibility, the control is in particular advantageous, or even a prerequisite, for the production of polypetides or RNA molecules in accordance with GMP standards. Typically, GMP materials are only approved if their production method is well-defined and typical parameters thereof are within an approved and pre-defined range of deviation. The inventive method thus controls and ensures that the production of a polypeptide or an untranslated RNA molecule in accordance with the invention is kept reproducible in such a manner that preferably no or only low fluctuation inbetween a certain threshold or pre-defined standard for specific parameters as a function of time occurs.

As used herein, the term "pre-defined standard" means that a standard is set as a function of time for parameters indicative for the reproducibility of the inventive method such as for the amount of polypeptide or untranslated RNA per cell being expressed by the recombinant host cells or for the number of nucleic acid molecule being in the recombinant host cells after its introduction herein. Such a pre-defined standard is typically determined by carrying out and repeating, respectively, the method of the invention, typically at least once, preferably at least twice, more preferably at least three times. The parameters being indicative for the reproducibility of the inventive method are standardized as a function of time, i.e. standardized typically at certain time points, by preferably calculating the arithmetic mean of the values for a given parameter at a specific time point. For example, to determine the amount of polypeptide or untranslated RNA or the number of nucleic acid molecule within the mentioned context aliquots of the medium supernatant and/or aliquots of the recombinant host cells are withdrawn at regular intervals, preferably daily during the production phase and are analyzed. The parameters indicative for the reproducibility of the inventive method analyzed during the production of the polypeptide or the untranslated RNA molecule are then compared with the pre-defined standard measured at the same time point. Upon determining the deviation of these values the production is continued if the deviation is within the pre-defined range of deviation, otherwise it is abandoned.

Thus, the invention provides a method for producing a polypeptide or untranslated RNA molecule conforming to FDA (US Federal Drug Administration) specification for Good Manufacturing Practice (GMP) production, comprising carrying out the production

steps according to one of the methods of the invention.

In a preferred embodiment, the nucleic acid molecule of the invention is an expression vector, preferably an expression vector capable of self-replication in a host cell, even more preferably an episomally replicating expression vector. Thus, the inventive nucleic acid molecules, expression systems and vector systems used in the method of the present invention allow for transient transfection, selection, and generation of stable cell populations, and very rapid production of polypeptides. As used herein, the term "vector" refers to a polynucleotide construct, typically a plasmid or a virus, used to transmit genetic material to a host cell. Preferably, the term "vector" as used herein refers to an agent such as a plasmid, and even more preferably to a circular plasmid. A vector as used herein may be composed of either DNA or RNA. Preferably, a vector as used herein is composed of DNA. The term "episomally replicating vector" or "episomal vector" is intended to mean a vector which is typically and very preferably not integrated into the genome of the host cell, but exists in parallel. An episomally replicating vector, as used herein, is replicated during the cell cycle and in the course of this replication the vector copies are distributed statistically in the resulting cells depending on the number of the copies present before and after cell division. Preferably, the episomally replicating vector replicates during S-phase of the cell cycle, preferably in the nucleus of the host cell. Moreover, the episomally replicating vector is replicated at least once, i.e. one or multiple times, in the nucleus of the host cell during S-phase of the cell cycle. In a very preferred embodiment, the episomally replicating vector is replicated once in the nucleus of the host cell during S-phase of the cell cycle.

In a further embodiment of the invention, methods are provided comprising introducing preferably inducible, potent and rapid expression systems that comprise one or more nucleic acid molecules into the at least one host cell. The nucleic acid molecules of the expression systems used in the method of the invention may be maintained at, or accumulated to, multiple copies in the nuclei of recombinant host cells, preferably of stable recombinant mammalian cells. In preferred embodiments, RNA self replication may lead to the accumulation of a high number of RNA molecules in the cytoplasm of the transfected cells and subsequent translation of the polypeptide of interest.

Thus, the nucleic acid molecule of the present invention comprises an expression system capable of self-replication in a host cell. Alternatively, the nucleic acid molecule of the present invention may comprise a vector system capable of replicating the nucleic

acid molecule in a host cell. The nucleic acid molecules as well as the expression systems and vector systems, respectively, of the present invention, thus, allow for semi-stable and highly efficient transfection and very rapid production of polypeptides eliminating the need of isolating stably transformed high producing single cell clones.

The nucleic acid molecules and expression systems of the invention comprise, inter alia, a first polynucleotide element capable of replicating the at least one nucleic acid molecule of the invention in a host cell. As used herein, the phrase "capable of replicating" means that the first polynucleotide element of the invention is capable of making an identical copy of a section of the nucleic acid molecule, preferably of the totality of the nucleic acid molecule of the invention, using existing nucleic acid molecules of the invention as a template for the synthesis of new nucleic acid strands.

In a preferred embodiment, the first polynucleotide element capable of replicating the at least one nucleic acid molecule of the invention comprises (a) a first polynucleotide sequence comprising an origin of replication; and (b) a second polynucleotide sequence encoding a replication initiation factor capable of recognizing the origin of replication. Thus, the replication initiation factor recognizes the origin of replication leading to replication of the nucleic acid molecule or expression vector, respectively, comprising the origin of replication. The first polynucleotide element of the invention may further comprise a third polynucleotide sequence comprising a plasmid maintenance factor.

The origin of replication of the invention refers to a DNA sequence that is recognized by a replication initiation factor or a DNA replicase leading to replication of a nucleic acid molecule or expression vector, respectively, containing the origin of replication. The expression "recognized by a replication initiation factor" is intended to mean that a replication initiation factor can physically interact with all or a portion of an origin of replication sequence, thereby causing or stimulating molecular mechanisms that ultimately cause all or a portion of the nucleic acid molecule comprising the origin of replication to be replicated. The origin of replication that is included in the methods of the invention may be derived from a prokaryotic organism, a eukaryotic organism (*e.g.*, a yeast, insect or mammal), and/or a virus. Preferably, the origin of replication is derived from a virus that allows for episomal replication. Preferably the origin of replication is derived from a DNA virus, more preferably from a DNA virus that allows for episomal replication. Most preferably, the origin of replication is derived from a DNA virus selected from the group consisting of Herpesviridae (e.g. Epstein-Barr virus, Herpes

simplex virus, Herpesvirus Saimiri, Murine Gammaherpesvirus 68, Human Cytomegalovirus, Mouse Cytomegalovirus, Pseudorabiesvirus), Papovaviridae (e.g. Simian Virus 40, Polyoma virus, human BK virus, Bovine Papiloma virus), Parvoviridae (Adeno-associated virus), Adenoviridae, and Hepadnaviridae. In a particularly preferred embodiment, the origin of replication is from Epstein-Barr virus (EBV), such as, *e.g.*, *ori*P, or is assembled from EBV sequences, and comprises EBNA-1 binding sites; examples for such EBV-type of functional origins are described in Aiyar et al., (1998) THE EMBO Journal, 17, pp. 6394-6403.

The replication initiation factor of the invention may be a protein or a DNA sequence encoding the open reading frame of such a protein or a coding sequence which can be separated by introns encoding for such a protein. Any functional variants of the replication initiation factor may also be used for the compositions and methods of the present invention. Such variants can be generated e.g. by substitutions, deletions, insertions or truncations of the amino acid of the replication initiation factor and its encoding DNA sequence, respectively. Methods for such are well known in the art and usually comprise specific site directed mutagenesis or generation of diversity by random mutagensis of which is then followed by selecting desired variants by means of functional assays. Routine methods employed for mutagenesis may be e.g. exposure to alkylating agents or UV irradiation, error-prone PCR or related gene shuffling PCR techniques and are usually performed in microorganisms (Miller, J., Experiments in Molecular Genetics, Cold Spring Harbor Laboratory 1972; Ling et al., 1997, Approaches to DNA Mutagenesis, Analytical biochemistry 254, 15 7-178; Cadwell et al., 1992, Randomization of genes by PCR mutagenesis in:PCR Methods, Cold Spring Harbor Laboratory Press 1992; Moore et al., 1997, Strategies for the in vitro evolution of protein function, J. Mol. Biol. 272, 336347). Preferred functional variants of the present invention are variants of the replication initiation factor with the same or increased replication activity as compared to the non-mutant replication initiation factor. Examples of such functional variants and methods how to test such functional variants are known in the art. For example, increased replication activity has been observed in EBNA1 mutants with a deletion in the ubiquitin-specific protease (USP-7) binding region (amino-acids 395-450) (*M. Holowaty et al., 2003, JBC, 29987-29994*). Another example for a functional variant of the replication initiation factor of the invention has been described by *Hung S.C. et al. (2000, PNAS, v98, 4, 1865-1870)* where the HMG-1 (high mobility group-1) amino-acids 1-90 could

substitute for EBNA-1 aminoacids 1-378 in mediating more efficient accumulation of replicated oriP plasmids, association with mitotic chromosomes, nuclear retention and episome maintenance. A further example for an improved functional variant of the replication initiation factor EBNA1, where amino acids 61-83 have been deleted, has been described by Wu H. et al (*2002, JV, 2480-2490).*

The replication initiation factor of the invention is derived from a prokaryotic organism, an eukaryotic organism (*e.g.*, yeasts, insects, and/or mammals) or a virus. Preferably, the replication initiation factor is derived from a virus that allows for episomal replication. Preferably, the replication initiation factor is derived from a DNA viruses, preferably a DNA virus selected from the group of Herpesviridae (e.g. Epstein-Barr virus, Herpes simplex virus, Herpesvirus Saimiri, Murine Gammaherpesvirus 68, Human Cytomegalovirus, Mouse Cytomegalovirus, Pseudorabiesvirus), Papovaviridae (e.g. Simian Virus 40, Polyoma virus, human BK virus, Bovine Papilloma virus), Parvoviridae (Adeno-associated virus), Adenoviridae, and Hepadnaviridae. Preferably, the replication initiation factor that is used with the present invention is a replication initiation factor from a DNA virus such as, *e.g.*, a Herpesvirus, Polyomavirus or a Papillomavirus which are episomally replicating viruses. More preferably, the replication initiation factor that is used with the invention is capable of operating as an episomal plasmid maintenance factor in the presence of the origin of replication. Most preferably, the replication initiation factor is a replication initiation factor from a DNA virus selected from the group consisting of Herpesviridae (e.g. EBNA-1 for Epstein-Barr virus (EBV)), Papovaviridae (e.g. the E1/E2 replication initiation factor for Papillomavirus such as BPV or HPV or e.g. large T-antigen replication initiation factor for Polyomaviridae such as SV40), Adenoviridae, and Hepadnaviridae.

In a particularly preferred embodiment, such replication initiation factor is from Epstein-Barr virus (EBV), such as, *e.g.*, the EBNA-1 protein, most preferably it is EBNA-1 protein or a functional variant thereof. The term 'functional variants' is used in its above meaning here. In the case of EBNA-1, this effect of enhanced episomal plasmid maintenance is achieved by the binding of the protein to the episome and, at the same time, to the chromosome, thereby leading to proper distribution of the episomes to the daughter cells.

The replication initiation factor may be constitutively expressed in the host cell of the invention, or alternatively, it may be inducibly expressed in the host cell of the

invention. Thus, the replication initiation factor preferably recognizes the origin of replication leading to replication of the nucleic acid molecule or vector, respectively, comprising the origin of replication.

The replication initiation factor that is used with the methods of the invention is, in certain embodiments, capable of recognizing the first origin of replication that is included within the nucleic acid molecule or expression system. Moreover, the replication initiation factor is capable of operating as a plasmid maintenance factor. As used herein, the term "plasmid maintenance factor" is intended to mean a factor, which supports the distribution of the episomes to the daughter cells upon cell division. In case of EBNA-1 this is achieved by binding of the protein to the episome and by the same time to the chromosome thereby leading to proper distribution of the episomes to the daughter cells. Thus, in certain embodiments, the replication initiation factor that is used with the invention is capable of operating as a plasmid maintenance factor. Alternatively, the first polynucleotide element of the invention may further comprise a third polynucleotide sequence comprising a plasmid maintenance factor. This means that not only the replication initiation factor is capable of operating as a plasmid maintenance factor, but that the first polynucleotide element may further to the origin of replication and the replication initiation factor comprise a third polynucleotide sequence which is capable of operating as a plasmid maintenance factor.

The origin of replication and the replication initiation factor that are included in, or used with, the compositions and methods of the invention may, in preferred embodiments, be derived from the same organism or the same virus. In an other embodiment, the origin of replication and the replication initiation factor may be derived from different organisms or viruses. Alternatively, the replication initiation factor may be derived from different organisms or viruses. As another alternative, the origin of replication may be derived from different organisms or viruses.

In a further embodiment, the first polynucleotide element of the invention comprises at least one element which attaches chromatin loop domains to the nuclear matrix fiber, forming the boundaries for these DNA loops (Gasser et al. (1989) Int. Rev. Cytol. 119:57 – 96), and thus has importance for structure and function. Preferably, this element is a scaffold/matrix attached region (S/MAR) element. The terms "scaffold/matrix attached region or scaffold/matrix attachment region (S/MAR)",

"S/MAR", "S/MAR element", "matrix attachment region (MAR)", and "scaffold attachment region (SAR)", are used interchangeably.

The S/MAR element used with the methods of the invention is preferably originating from the 5'region of the human interferon beta gene (SEQ ID NO: 8). Further S/MAR elements within the scope of the invention are any eukaryotic DNA elements which are involved in the generation of the loop structure found in chromatin. Such elements are typically A/T rich, in the range of 70% and show high affinity binding to the nuclear matrix. Methods for the identification of such S/MAR elements have been described (Sumer et al. Genome Res. 2003 Jul;13(7):1737-43). A comprehensive listing of S/MAR elements is found in the S/MAR t database, currently under http://transfac.gbf.de/SMARtDB/ and has been described by Liebich, I., Bode, J., Frisch,M. and Wingender, E.: S/MARt DB: A database on scaffold / matrix attached regions. Nucleic Acids Res. 30, 312-374 (2002).

Within the scope of the present invention are S/MAR elements that originate from mammals selected from the group consisting without limitation of human, simian, sheep and rodents, and herein typically mouse, rat, rabbit or hamster. Other S/MAR elements that can be used for the present invention are of chicken origin, of yeast, Caenorhabditis elegans, Drosophila melanogaster or virus origin and herein typically of EBV, SV40 or papilloma virus origin. In yet another embodiment, the S/MAR element is of plant origin, typically of tomato, Arabidopsis thaliana, tobacco, rice, pea, maize, potato, soybean, petunia hybrida, or sorghum bicolour origin. Preferred S/MAR elements useful in the present invention are of mammalian origin and herein in particular of human origin, most preferably originating from the 5'region of the human interferon beta gene.

Preferred nucleic acid molecules or expression vectors, respectively, that are used in the methods of the invention, comprising an origin of replication, preferably OriP, a replication initiation factor, preferably EBNA-1, and a S/MAR element have been described in the US provisional application entitled "Gene Expression System" filed by the present assignee on November 26, 2003, the disclosure of which is hereby incorporated by reference in its entirety.

The nucleic acid molecules of the present invention further comprise at least one second polynucleotide element comprising at least one nucleotide sequence selected from the group consisting of (i) at least one open reading frame encoding a polypeptide of interest; (ii) a nucleotide sequence complementary to all or part of the open reading frame

of (i); and (iii) a nucleotide sequence encoding an untranslated RNA molecule or complement thereof. The second nucleotide sequence, in certain embodiments, may further comprise other nucleotide sequence of interest, as well as other elements.

In certain embodiments of the present invention, the second polynucleotide element may encode an RNA molecule, the RNA molecule comprising (a) at least one cis acting sequence element; (b) a first nucleotide sequence comprising a first open reading frame, the first open reading frame having a nucleotide sequence encoding an RNA dependent RNA polymerase; and (c) at least one second nucleotide sequence selected from the group consisting of (i) an open reading frame encoding a polypeptide of interest; (ii) a nucleotide sequence complementary to all or part of the second open reading frame of (i); and (iii) a nucleotide sequence encoding an untranslated RNA molecule or complement thereof; wherein the second nucleotide sequence is operably linked to a promoter which is recognized by the RNA-dependent RNA polymerase. The first nucleotide sequence may, in certain embodiments, also comprise 5' and 3' cis acting sequences, and junction sequences containing a subgenomic promoter. The replicase protein coding sequences, the 5' and 3' cis acting sequences (when present), and the junction sequences containing the subgenomic promoter will normally be derived from a virus, preferably from an alphavirus, most preferably from Sindbis virus.

In such embodiments, nucleic acid molecules and expression systems of the present invention are based on the vectors used in the expression systems described in WO 99/50432 and in PCT/EP03/09291, the disclosure of which is hereby incorporated by reference in its entirety. These vectors are constitutively transcribed in host cells to produce mRNA molecules having two open reading frames, e.g., a first open reading frame which encodes a replicase, and a second open reading frame which encodes a polypeptide of interest. These open reading frames, which may or may not be produced from the same nucleic acid molecule, encode a temperature-sensitive replicase and a heterologous gene of interest. The term "heterologous gene of interest" refers to any amino acid or RNA sequence encoded by a heterologous DNA sequence contained in a vector of the invention. Heterologous nucleotide sequences can encode polypeptides or RNA molecules normally expressed in the cell type in which they are present or molecules not normally expressed therein (e.g., Sindbis structural proteins). The first open reading frame having a nucleotide sequence encoding an RNA-dependent RNA polymerase (replicase) is translated to produce an RNA dependent RNA polymerase. The

second open reading frame, encoding all or part of one or more polypeptides of interest, is not translated until after at least one RNA dependent RNA replication event. The RNA dependent RNA replication event is intended to mean a process which results in the formation of an RNA molecule using an RNA molecule as a template.

As used herein, the phrase "cis-acting" sequence refers to nucleic acid sequences to which a replicase binds to catalyze the RNA dependent replication of RNA molecules. These replication events result in the replication of the full-length and partial RNA molecules and, thus, the alphavirus subgenomic promoter is also a "cis-acting" sequence. Cis-acting sequences may be located at or near the 5' end, 3' end, or both ends of a nucleic acid molecule, as well as internally.

As used herein, the phrase "RNA-dependent RNA polymerase" refers to a polymerase which catalyzes the production of an RNA molecule from another RNA molecule. This term is used herein synonymously with the term "replicase."

The RNA-dependent RNA polymerase that is included in, or used with, certain methods of the invention may, in certain embodiments, be selected from the group consisting of: (a) a temperature-sensitive RNA-dependent RNA polymerase, (b) a non-cytopathic RNA-dependent RNA polymerase, and (c) a temperature-sensitive, non-cytopathic RNA-dependent RNA polymerase.

Thus, in certain embodiments, RNA self-replication is preferably inducible and triggered by a temperature shift. The high production levels, in certain embodiments, are achieved by the use of a temperature-sensitive RNA-dependent RNA polymerase (i.e., a replicase) which only replicates RNA molecules to form new RNA molecules at permissive temperatures. Moreover, if desired, the amount of RNA or polypeptide can be tightly regulated by the use of a temperature-sensitive RNA-dependent RNA polymerase. This facilitates, in particular, production of growth-inhibitory or toxic polypeptides or RNA-species. When the RNA-dependent RNA polymerase that is encoded by the first open reading frame of the nucleic acid molecules of the invention is a temperature-sensitive RNA-dependent RNA polymerase, the RNA-dependent RNA polymerase may, e.g., have replicase activity at temperatures below 34°C and low or undetectable replicase activity at 34°C or above. In certain embodiments, the temperature sensitive RNA-dependent RNA polymerase will be one that has replicase activity at 34°C which is at least five fold lower than the replicase activity exhibited at 29°C.

The RNA-dependent RNA polymerase that is included in, or used with, the compositions and methods of the invention may, in certain embodiments, be of viral origin, preferably of alphaviral origin, more preferably derived from a Sindbis virus, a Semliki Forest virus or an Aura virus. As used herein, the term "alphavirus" refers to any of the RNA viruses included within the genus Alphavirus. Descriptions of the members of this genus are contained in Strauss and Strauss, Microbiol. Rev., 58:491-562 (1994). Examples of alphaviruses are selected from the group comprising Aura virus, Bebaru virus, Cabassou virus, Chikungunya virus, Easter equine encephalomyelitis virus, Fort morgan virus, Getah virus, Kyzylagach virus, Mayoaro virus, Middleburg virus, Mucambo virus, Ndumu virus, Pixuna virus, Tonate virus, Triniti virus, Una virus, Western equine encephalomyelitis virus, Whataroa virus, Sindbis virus (SIN), Semliki forest virus (SFV), Venezuelan equine encephalomyelitis virus (VEE), and Ross River virus.

Thus, in one certain embodiment, the present invention provides a hybrid vector and vector system combining the advantages of EBV episomal replication with the advantages of RNA virus replication, wherein the replication takes place entirely in the cytoplasm. This type of hybrid vector would be characterized by the maintenance of multiple copies of DNA due to polynucleotide replication, typically plasmid replication, followed by the accumulation of full-length transcripts containing the viral replicon. These viral full-length transcripts are then further exponentially amplified in the cytoplasm by the viral replicase leading to the accumulation of RNA virus replicons and protein synthesis. The system is extremely valuable for production of toxic polypeptides if the RNA virus replication is controllable or inducible.

In a specific embodiment, the invention includes the combination of a Herpesvirus mini-replicon unit, i.e. the cis-acting replication origin OriP (Origin of replication P) and the cis- or trans-acting gene product – EBNA-1 (Epstein-Barr virus nuclear antigen 1, a replication triggering factor) with the tightly regulated temperature inducible alphaviral expression system, as disclosed in WO 99/504332. Due to the presence of OriP, EBNA-1 and pCytTS, the pCytTS-OPE vector (the construction of which is described in Example 2 of the PCT application with the number PCT/EP03/09291) is maintained episomally (extrachromosomally) in the form of several DNA copies in the cell nuclei. These multiple DNA copies may be transcribed from a promoter, typically a CMV or RSV promoter, into CytTs RNA-replicons, but they remain inactive unless cells are being

shifted to a certain temperature. Only after temperature induction replicon replication takes place, followed by RNA accumulation and translation in the cell cytoplasm and production of the polypeptide of interest. The term pCytTs may include any temperature sensitive, non-cytopathic, inducible alphaviral expression system containing any functional promoter to drive the transcription of mRNA from the nucleic acid molecule of the invention. Examples for such expression systems or vectors may be, e.g., the vector pCytTs containing a RSV promoter or the vector pCytTs2.1 containing a CMV promoter. When combinatorial constructs with pCytTs are mentioned in the Examples, they generally contain a CMV promoter, even if the abbreviation of the combinatorial constructs does – for the sake of simplicity - only refer to pCytTs and not to pCytTs2.1.

When sequences encoding the polypeptide of interest and sequences encoding an RNA-dependent RNA polymerase are present either on the same nucleic acid molecule or in the same vector (i.e., in a single vector system), a region will preferably be present 5' to the second open reading frame which inhibits translation of this open reading frame.

The nucleic acid molecules of the invention, in certain embodiments, comprise a 5' promoter which is capable of initiating transcription in vivo, 5' and/or 3' sequences enabling replication of the RNA molecule (cis-acting sequence elements), and a subgenomic promoter 5' to the gene of interest, as well as a sequence of interest which is translatable only after one or more RNA dependent RNA replication events. These RNA dependent RNA replication events are catalyzed by a, preferably regulatable, RNA dependent RNA polymerase which may be encoded by the same mRNA molecule that is produced by transcription of the DNA vector or by a different mRNA molecule.

The second polynucleotide element of the invention that is included within the nucleic acid molecules and expression systems of the invention may be, e.g., an open reading frame encoding a polypeptide of interest. Such an open reading frame may, alternatively or interchangeably, be referred to as a nucleotide sequence or polypeptide of interest. A wide variety of nucleotide sequences of interest can be expressed by the nucleic acid molecules and expression systems of the invention. These genes of interest include, but are not limited to, sequences encoding adhesion molecules, antibodies, clotting factors, receptors, regulatory proteins, structural proteins, transcription factors, transport proteins, structural proteins, toxins, enzymes, prodrug converting enzymes, antigens which stimulate immune responses, single chain antibodies, polypeptides which stimulate or inhibit immune responses, tumor necrosis factors, light and/or heavy chain of

an antibody, lymphokines (e.g. beta-interferon), cytokines, and various proteins with therapeutic applications (*e.g.*, growth hormones and regulatory factors). Hematopoiesis is regulated by lymphokines and cytokines which stimulate the proliferation and/or differentiation of various hemopoietic cells. Representative examples of cytokines and lymphokines include interleukin-1 (IL-1), interleukin-2 (IL-2), interleukin-3 (IL-3), interleukin-4 (IL-4), interleukin-5 (IL-5), interleukin-6 (IL-6), interleukin-7 (IL-7), interleukin-8 (IL-8), interleukin-9 (IL-9), interleukin-10 (IL-10), interleukin-11 (IL-11), interleukin-12 (IL-12), interleukin-13 (IL-13), interleukin-14 (IL-14), interleukin-15 (IL-15), interleukin-16 (IL-16), interleukin-17 (IL-17), granulocyte colony stimulating factor (G-CSF), granulocyte-macrophage colony stimulating factor (GM-CSF), macrophage colony stimulating factor (M-CSF), and interferons. The second open reading frame may further encode resistin or leptin.

In a preferred embodiment, the nucleic acid molecule or expression system, respectively, of the invention comprises two polypeptides of interest for example for the expression of antibodies. Alternatively, the nucleic acid molecule or expression vector, respectively, may comprise several polypeptides of interest like for example for the expression of multi subunit proteins such as ion channels. The individual polypeptides will be expressed by different promoters or by one promoter in the presence of internal ribosomal entry sites. The different promoters can be of the same or different origins. In a preferred embodiment, the nucleic acid molecule or expression vector, respectively, comprises a first open reading frame encoding the light chain of an antibody, and a second open reading frame encoding the heavy chain of an antibody. Preferably, the light chain and heavy chain are of the same antibody. Thus, in certain embodiments, the second polynucleotide element of the invention may encode an antibody. The complete antibody may be expressed from two different promoters, for example light chain from a first CMV promoter and the heavy chain from a second CMV promoter. Additionally single chain antibodies may be expressed.

In certain embodiments, the second polynucleotide element of the invention may encode secreted enzymes (e.g., secreted alkaline phosphatase), cytoplasmic enzymes (e.g., green fluorescent protein), or any number of other proteins with therapeutic applications (e.g., human insulin, human coagulation Factor VIII).

The nucleic acid molecules and expression systems used with the methods of the invention can also comprise a second polynucleotide element that can be used to express

heterologous sequence encoding cytotoxic polypeptides. Cytotoxic polypeptides act to directly or indirectly inhibit cellular growth or metabolism. Representative examples of toxins include Shigella toxin, ricin, Diphtheria toxin, Cholera toxin, Pseudomonas exotoxin A, and Herpes simplex virus thymidine kinase (HSVTK). Within other embodiments of this invention, the heterologous sequence encodes a prodrug converting enzyme. A prodrug converting enzyme activates a compound with little or no cytotoxicity into a toxic product. Representative examples are HSVTK, alkaline phosphatase, guanine phosphoribosyl transferase, and penicillin V amidase. Examples of both cytotoxic polypeptides and prodrug converting enzymes are discussed in numerous sources including WO 97/38087, EP 0716148, and WO 96/17072. In addition, a vast array of signaling molecules and membrane proteins are toxic if expressed at high levels. All these molecules may be suitable for expression using the system of the present invention.

The nucleic acid molecules and expression systems used with the methods of the invention can also comprise a second polynucleotide element that can be used to express virtually any polypeptide, including ones which have not as yet been identified but are encoded by nucleotide sequences contained in, for example, cDNA libraries or host cell chromosomes. Examples of such polypeptides include secreted proteins and proteins from various cellular compartments. Heterologous sequences expressed by the vectors of the invention can encode polypeptides and RNA molecules from non human species (e.g., other mammals, plants, fungi, bacteria or viruses). These heterologous sequences may further encode viral membrane proteins (e.g., HIV gp160) or viral polyproteins (e.g., Sindbis structural proteins).

Nucleotide sequences may be added to the nucleic acid molecules and vectors of the invention which result in the production of a fusion protein. For example, such sequences can encode amino acid sequences which are fused to a protein encoded by a gene of interest and confer one or more functional characteristics upon the expression product. These amino acid sequences include sequences which will target the gene product for export from the cell (e.g., a secretory sequence) or to a subcellular compartment (e.g., the nucleus). Such amino acid sequences further include sequences which facilitate purification (e.g., a six His "tag"). Depending on the amino acid sequence and the function imparted by the fused sequence, the added amino acid sequences may or may not be cleaved from the translation product.

Fusion proteins also include proteins which have domains or regions derived from various different proteins. Examples of such a fusion protein are those containing domain II of Pseudomonas exotoxin, a domain or amino acid sequence which has binding affinity for a cell surface receptor associated with a particular cell type, and another amino acid sequence with a preselected biological activity. Domain II of Pseudomonas exotoxin will translocate across cell membranes. Using this system, fusion proteins can be designed which will bind to specific cells types, will translocate across the cytoplasmic membranes of these cells, and will catalyze predetermined intracellular biological reactions. A system of this type is described in Pastan et al., U.S. Patent No. 5,705,163. Methods for identifying amino acid sequences, which bind to specific cell types are described in Wu, A., Nature Biotech. 14:429-431 (1996).

The nucleic acid molecules and expression systems of the invention may, in certain embodiments, further comprise genetic elements which confer additional functional characteristics such as selection markers.

Thus, the nucleic acid molecules of the invention may further comprise a selection marker. As used herein, the phrase "selection marker" refers to a marker gene present on a vector of the invention whose expression allows one to identify cells that have been transformed or transfected with a vector containing the marker gene. The selection marker may facilitate the cloning and amplification of the vector sequences in prokaryotic and eukaryotic organisms, thus the selection markers of the invention are useful for selection in eukaryotic host cell, as well as for positive selection in bacterial host cells. Any marker known in the art may be used, including without limitation, e.g., markers that confer resistance to a compound or class of compounds, such as antibiotics, so that only cells expressing nucleic acid molecules comprising a selection marker may survive when grown in the presence of e.g. antibiotics. Selection markers for eukaryotic cells may be used that confer resistance to e.g. puromycin, hygromycin, gpt, neomycin, zeocin, ouabain, blasticidin, bleomycin, or markers such as DHRF, hisD, glutamine synthetase, and trpB. Preferably, the selection marker comprises resistance to puromycin. In another preferred embodiment, the selection marker is glutamine synthetase.

The vectors of the present invention may also contain a selection marker for positive selection in bacterial host cells, e.g. ampicillin, tetracycline, kanamycin, chloramphenicol, or e.g. sequences which result in high copy number amplification, such as e.g. an E. coli origin of replication (i.e., the ColE1 sequence contained in the pCytTs

vector). A considerable number of sequences encoding additional selection markers and origins of replication are known in the art (see, e.g., Sambrook, J. et al., eds., MOLECULAR CLONING, A LABORATORY MANUAL, 2nd. edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989); Ausubel, F. et al., eds., CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John H. Wiley & Sons, Inc. (1997)).

The nucleic acid molecules and expression systems of the invention can be also used to express more than one gene of interest. For example, recombinant host cells can be transfected with more than one nucleic acid molecule of the invention wherein one nucleic acid molecule encodes both the replication initiation factor and a polypeptide of interest and additional nucleic acid molecules may encode additional polypeptides of interest.

The nucleic acid molecule of the invention, in further embodiments, comprises single and multiple vector systems for producing a polypeptide or untranslated RNA molecule. In a single vector system, the first, second, and third polynucleotide element, as well as the first, second, and third polynucleotide sequence of the invention are all present on the same nucleic acid molecule. In a multiple vector system, the first and second polynucleotide elements of the invention are present on one or more separate nucleic acid molecules. Alternatively, the first and second polynucleotide element as well as the first, second and third polynucleotide sequences of the invention are each present on one or more separate nucleic acid molecules. In certain embodiments, the first and second polynucleotide sequences will be on a single nucleic acid molecule while the second polynucleotide element is on a different nucleic acid molecule. In certain other embodiments, the first polynucleotide sequence and the second polynucleotide element will be on a single nucleic acid molecule, while the second polynucleotide sequence is on a different nucleic acid molecule. In certain other embodiments, the second polynucleotide sequence and the second polynucleotide element will be on a single nucleic acid molecule, while the first polynucleotide sequence is on a different nucleic acid molecule. Alternatively, the polypeptide of interest and the nucleotide sequence encoding an RNA-dependent RNA polymerase of the invention will be on one or more seperate nucleic acid molecules. In such a multiple vector system, at least one, preferably some or more preferably all vectors contain an origin of replication. In addition, all vectors may contain a third polynucleotide element of the invention comprising a

selection marker. In preferred embodiments of the invention, the replication initiation factor and the origin of replication recognizing the replication initiation factor may be on the same nucleic acid molecule and plasmid, respectively, as it is the RNA-replicase or on a separate nucleic acid molecule and plasmid, respectively.

In the multiple vector system, a 5' sequence can inhibit translation by having one or more short open reading frames with associated stop codons which lead to the detachment of ribosomes. Similarly, any sequence which inhibits the traveling or binding of ribosomes to the sequence of interest can be used as a 5' sequence which inhibits translation (Voet and Voet, BIOCHEMISTRY, John Wiley & Sons, Inc. (1990)).

Another method for preventing translation of nucleotide sequences in most biological systems involves the insertion of the sequence in an antisense direction. This method of inhibiting translation is based on the principle that the presence of antisense RNA results in the downregulation of the sense RNA therefore reducing translation.

The second polynucleotide element that is included in, or used with, the methods of the invention may encode, in certain embodiments, an untranslated RNA molecule or complement thereof. As used herein, the phrase "untranslated RNA" refers to an RNA sequence or molecule which does not contain an open reading frame or encodes an open reading frame, or portion thereof, but in a format in which an amino acid sequence will not be produced (e.g., no initiation codon is present). For example, RNA molecules directly produced by transcription of a DNA sequence of the invention may encode RNA sequences which are neither translated nor present in an antisense orientation. Exemplary untranslated RNA molecules include, e.g., an antisense RNA molecule, a tRNA molecule, an rRNA molecule, RNase P targeted sequences which induce gene down regulation, and ribozymes. Smith S. et al. (J. Virol. 71:9713 9721 (1997)) describes alphaviral vectors used to express ribozyme sequences. Antisense RNA may be untranslated but, in some instances antisense sequences can be converted to a translatable sense strand from which a polypeptide is produced. Other useful untranslated RNA molecules are known in the art.

The nucleic acid molecules and vector systems of the invention may contain a promoter sequence which drives transcription to produce mRNA molecules having coding sequences of the open reading frames of the invention. In one embodiment, the open reading frame preferably encodes one or more polypeptides of interest. In addition, the vectors contain an origin of replication, a replication initiation factor recognizing the origin of replication, and a eukaryotic selection marker. In multiple vector systems, each

vector may contain a promoter sequence which drives transcription to produce mRNA molecules having coding sequences of the open reading frames of the invention, an origin of replication and a eukaryotic selection marker.

While any functional promoter can be used to drive the transcription of mRNA from the nucleic acid molecule of the invention, the promoter is preferably a constitutive RNA polymerase II promoter (e.g., Rous Sarcoma Virus (RSV), cytomegalovirus (CMV), simian virus 40 (SV40), myeloproliferative sarcoma virus (MPSV), glucocorticoid, metallothionein, Herpes simplex virus thymidine kinase (HSVTK), human immune deficiency (HIV), EF1α (elongation factor 1alpha (Mizushima et al., 1990, Nuc. Acid Research, Vol.18, 5322), mouse mammary tumor virus (MMTV), human polyomavirus BK (BKV), or Moloney murine leukemia virus (MuLV) promoter). Additional promoters suitable for use in the practice of the present invention are known in the art (see, e.g., Lee, A. et al., Mol. Cells. 7:495 501 (1997)).

Further, one or more of these DNA vectors can be designed to stably integrate into the host cell genome. When expression of a gene of interest is desired in a cell type containing one or more stably integrated DNA molecules of the invention, expression of the gene of interest will require the introduction of nucleic acid molecules (DNA or RNA) encoding the components of the system into the cells not present in the integrated molecule(s).

The invention is also directed to the production of polypeptides or RNA molecules of interest using mammalian cells grown in serum free or protein free culture media. For example, by long term culture under conditions restricting serum access or selecting for suspension growth, CHO cell lines are selected which are able to grow in serum-free medium and/or in suspension (Zang. M. et al., Bio/Technology 13:389 (1995)).

Thus, the invention includes methods of producing recombinant host cells. For example, the methods of the invention comprise introducing one or more nucleic acid molecules or expression systems described herein into at least one host cell. According to a certain aspect, the at least one host cell, and host cell line respectively, provided for the method of the present invention is useful for the propagation of replication-competent expression vectors and growth under serum-free conditions. This includes the production of a recombinant cell line which makes the cell line competent for the replication of the expression vector. Thus, at least one recombinant host cell of the present invention and in vitro cell cultures comprise at least one nucleic acid molecule or at least one vector

system of the invention. The at least one host cell may be one or more host cell, preferably a host cell line, in particular a culture of a particular type of cell that can be reproduced, preferably, indefinitely. The host cell of the invention may be a prokaryotic or eukaryotic host cell. Alternatively, more than one host cell line may be provided for the method of the present invention. In one embodiment, one, some or all of these recombinant host cells contain one or more nucleic acid molecules of the present invention. The first polynucleotide element of the invention or parts thereof may be stably integrated into the genome of the one, some or all host cells. Alternatively, the replication initiation factor of the present invention may be stably integrated into the genome of the one, some or all host cells. A variety of different recombinant host cells can be produced which contain the nucleic acid molecules and expression systems of the invention. Alphaviruses, for example, are known to have a wide host range. Sindbis virus, for example, infects cultured mammalian, reptilian, and amphibian cells, as well as some insect cells (Clark, H., J. Natl. Cancer Inst. 51:645 (1973); Leake, C., J. Gen. Virol. 35:335 (1977); Stollar, V. in THE TOGAVIRUSES, R.W. Schlesinger, Ed., Academic Press, (1980), pp.583 621).

Thus, numerous host cells can be used in the practice of the invention. Representative host cells that may be used with the invention include, but are not limited to, bacterial cells, yeast cells, plant cells and animal cells. Preferred bacterial host cells include Escherichia spp. cells (particularly *E. coli* cells and most particularly E. coli strains DH10B, Stbl2, DH5, DB3, DB3.1, DB4 and DB5), Bacillus spp. cells (particularly B. subtilis and B. megaterium cells), Streptomyces spp. cells, Erwinia spp. cells, Klebsiella spp. cells, Serratia spp. cells (particularly S. marcessans cells), Pseudomonas spp. cells (particularly P. aeruginosa cells), and Salmonella spp. cells (particularly S. typhimurium and S. typhi cells). Preferred animal host cells include insect cells (most particularly Drosophila melanogaster cells, Spodoptera frugiperda Sf9 and Sf21 cells and Trichoplusa High-Five cells), nematode cells (particularly C. elegans cells), avian cells, amphibian cells (particularly Xenopus laevis cells), reptilian cells, and mammalian cells (most particularly human, simian, canine, rodent, bovine, or sheep cells, e.g. NIH3T3, CHO (Chinese hamster ovary cell), COS, VERO, BHK, HEK, and other rodent or human cells). Preferred yeast host cells include Saccharomyces cerevisiae cells and Pichia pastoris cells. In another embodiment, BHK, COS, Vero, HeLa and CHO cells are further preferred since they are particularly suitable for the production of heterologous

polypeptides because they have the potential to glycosylate heterologous proteins in a manner similar to human cells (Watson, E. *et al., Glycobiology 4*:227, (1994)) and can be selected (Zang, M. *et al., Bio/Technology 13*:389 (1995)) or genetically engineered (Renner W. *et al., Biotech. Bioeng. 47*:476 (1995); Lee K. *et al. Biotech. Bioeng. 50*:336 (1996)) to grow in serum-free medium, as well as in suspension.

In one particularly preferred embodiment of the invention, the at least one host cell is a rodent cell, more preferably a hamster cell, most preferably a CHO or a CHO derived cell. CHO derived cells can be e.g. cells obtained from originating CHO cell lines by genetic engineering, cell fusion, extended passaging and medium-adaptation process and the like. Preferred CHO cells and CHO derived cells suitable for the production of polypeptides include without limitation CHO-K1 (e.g. ACC-110 deposited at DSMZ, German depository acknowledged under Budapest Treaty), CHO-DUKX , CHO-DXB11, CHO-dhfr (e.g. ACC-126 deposited at DSMZ, German depository acknowledged under Budapest Treaty), CHO-F3B4, CHO-GD3, CHO-SSF3, B13-24, 35.6, 5/9m alpha 3-18 (e.g. CRL-10154 deposited at ATCC, U.S.A. depository acknowledged under Budapest Treaty), 6E6 (e.g. CRL-11398 deposited at ATCC, U.S.A. depository acknowledged under Budapest Treaty), AA8 cells, and CHO derived cells expressing the genomically integrated EBNA1 protein.

In another preferred embodiment of the present invention, the host cell is an animal cell, preferably a mammalian cell, more preferably a human cell, most preferably a 293 derived cell. Preferred cells for the production of polypeptides according to the invention are 293 HEK (human embryonic kidney) cells or 293 EBNA1 cells. Other 293 derived cells known in the art can be used for the present invention, such as for example 293 derived cells adapted to serum-free growth conditions.

In a preferred embodiment, host cells are adapted to serum-free or protein-free growth prior to transfection with the nucleic acid molecule of the invention. Non-limiting examples of host cells for use in the present invention include HEK 293 cells (ATCC, Cat. No. CRL-1573), CV-1EBNA cells (ATCC, Cat. No. CRL-10478), Hela (ATCC, Cat. No. CCL-2), Raji cells (ATCC, Cat. No. CCL-86), 293 EBNA (Invitrogen, Cat. No. R62007), CV1 cells (ATCC, Cat. No. CCL-70) and 143B cells (ATCC, Cat. No. CRL-8303). Alternatively, for performing the method of the invention host cells are provided that already are adapted to serum-free or protein-free growth. Numerous cell are available commercially that are adapted to serum-free or protein-free growth (e.g. 293-F and 293-H

cells, human embryonic kidney cells adapted to serum-free suspension culture, Invitrogen, Cat. No. 11625, 11631, respectively; CHO-S, Chinese Hamster Ovary, suspension adapted, Invitrogen, Cat. No. 11619; BHK21C13-2P, Hamster Syrian kidney, adapted to growth in suspension, ECACC, Cat. No. 84111301; HeLa S3, human negroid cervix carcinoma, ECACC, Cat. No. 87110901, HeLa/SF, human negroid cervix carcinoma;) .

The nucleic acid molecules and/or expression systems of the invention may be introduced into host cells using well known techniques of infection, transduction, electroporation, transfection, and transformation. Exemplary methods include DEAE-dextran mediated transfection, transient transfection, microinjection, cationic lipid-mediated transfection, scrape loading and ballistic introduction. Methods for the introduction of exogenous DNA sequences into host cells are discussed in Felgner, P. et al., U.S. Patent No. 5,580,859. The nucleic acid molecules and/or vectors of the invention may be introduced alone or in conjunction with other nucleic acid molecules and/or vectors and/or proteins, peptides or RNAs. Alternatively, the nucleic acid molecules and/or expression systems of the invention may be introduced into host cells as a precipitate, such as a calcium phosphate precipitate, or in a complex with a lipid. Electroporation also may be used to introduce the nucleic acid molecules and/or expression systems of the invention into a host. Likewise, such molecules may be introduced into chemically competent cells such as E. coli. If the vector is a virus, it may be packaged in vitro or introduced into a packaging cell and the packaged virus may be transduced into cells. Hence, a wide variety of techniques suitable for introducing the nucleic acid molecules and/or vectors of the invention into host cells are well known and routine to those of skill in the art. Such techniques are reviewed at length, for example, in Sambrook, J., et al., Molecular Cloning, a Laboratory Manual, 2nd Ed., Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, pp. 16.30 16.55 (1989), Ausubel, F. et al., eds., CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John H. Wiley & Sons, Inc. (1997), Chapter 16), Watson, J.D., et al., Recombinant DNA, 2nd Ed., New York: W.H. Freeman and Co., pp. 213 234 (1992), and Winnacker, E. L., From Genes to Clones, New York: VCH Publishers (1987), which are illustrative of the many laboratory manuals that detail these techniques and which are incorporated by reference herein in their entireties for their relevant disclosures. Preferably, the introduction of a nucleid acid molecule into a host cell to produce a recombinant host cell is performed by transfection. Transfection techniques well known in the art can be used, e.g. selecting without

limitation from techniques of transfection using polyethyleneimine (PEI) (Boussif et al., 1995, PNAS, 92, 7297-7301), calcium phosphate copreicipitation method (Graham et al., 1973, Virology, 52, 456-467), continous flow electroporation (Parham et al., Cytotechnology, 28, 147-155), or cationic lipid reagents (Felgner et al., 1987, Proc. Natl. Acad. Sci. 84, 7413).

In an exemplary embodiment of the invention, nucleic acid molecules are provided that are constructed by introducing defined Epstein-Barr virus (EBV) sequences into vectors such as, e.g., those included within the pCep or pCytTs system. Other vectors may include but are not limited to e.g. pBS, pPUR, pcDNA. The resulting exemplary nucleic acid molecules of the invention are capable of replicating as non-integrated autonomous episomal molecules in the transformed host cells.

The defined sequences, which may be introduced into vectors, comprise, e.g., either one or two elements of Epstein-Barr virus, (OriP alone or together with EBNA-1 gene) which permit plasmid maintenance. The origin of replication OriP is a cis-acting sequence and needs to be inserted into the plasmid vector sequence. Plasmids containing this origin of replication are able to be maintained in cells expressing the replication initiation factor which recognizes the origin of replication; one of these factors is EBNA-1. EBNA-1 gene function can be provided in cis by introducing the sequence into the same plasmid or can be provided in trans by co-transfection with a second replicating or non-replicating plasmid as well as by providing it from stably transduced cell lines expressing EBNA-1 gene from an integrated copy. Furthermore, in certain embodiments, more than one copy of the sequence expressing the replication initiation factor, e.g. EBNA-1, is provided. For example, EBNA-1 gene function may be provided on one or more plasmids, wherein none, some or all of those plasmids may additionally comprise the origin of replication, and which plasmids are used for transfecting cell lines, preferably cell lines that have the sequence of EBNA-1 stably integrated within its genome. Presence of both OriP and EBNA1 sequences in the same plasmid leaves replication less dependent from the host cell type. (Reisman, D. et al., 1985, MCB, 5, 1822-1832, Yates, J.L. et al., 1985, Nature, 313, 812-815, U.S.Pat.No 4,686,186). U.S.Pat.No 4,686,186 describes the transfection of cells with a single plasmid containing the EBV OriP, the EBNA-1 gene and a gene encoding a protein of interest.

Transfection of cell lines that express EBNA 1 can be advantageous since the ability of such cells to stably maintain an episomal construct can be enhanced by several

orders of magnitude, and stable cell lines can be generated in as little as two to three weeks. For example, HEK cells that stably express EBNA 1 have been transformed with plasmids containing the EBV origin of replication, and the gene encoding CRHR1 (corticotropin releasing hormone receptor subtype I). The resulting cell lines have been found to stably express high levels of CRHR1. (Horlick et al., Prot. Exp. And. Purific. 9:301-308, 1997.)

Although such cell lines have the advantage of stable long-term expression of the replication initiation factor and durable support of replication and maintenance of OriP containing plasmids, there are not many EBNA-1 expressing cell lines commercially available (ATCC : 293HEK-EBNA1 and CV1-EBNA1). Alternatively, plasmids that already carry the EBNA-1 gene and the gene of interest in cis on the same episome are used to transfect cells and commercial vectors such as pCEP4 (Invitrogen) are available. However, current vectors designed for constitutive expression of polypeptides carrying OriP and EBNA-1 on the episomal construct in cis and wherein the expression is neither regulatable nor inducible, are not applicable in cases where cell-toxic polypeptides need to be expressed.

Thus, in certain exemplary embodiments, the invention provides methods which allow for the production of high amounts of specific RNA molecules produced in host cells upon their transfection with polynucleotides. This high production results, e.g., from the use of a self-replicating episomal DNA plasmid, which may be, in one embodiment, combined with an RNA-dependent RNA polymerase, which amplifies the mRNA of interest. Exemplary nucleic acid molecules of the invention may comprise without limitation polynucleotides, which comprise the nucleotide sequence of pCytTs-OriP, the sequence of pCytTs-OPE, or SEQ ID NO: 1 (pCEPpu), and thus polynucleotides which contain the origin of replication, oriP, and the replication initiation factor EBNA-1 (see Example 1).

In a further aspect, the invention provides a method for producing a polypeptide or untranslated RNA molecule conforming to FDA specification for GMP production, comprising carrying out the production steps according to one of the methods of the invention.

The invention also provides methods for regulating the expression of a polypeptide or an untranslated RNA molecule. The methods of the invention may comprise, e.g., (a) introducing one or more nucleic acid molecules or expression systems of the invention

into at least one host cell to produce a recombinant host cell, wherein the nucleic acid molecules or expression systems comprise a first open reading frame having a nucleotide sequence encoding a temperature-sensitive RNA-dependent RNA polymerase, (b) growing the recombinant host cell under suitable culture conditions, (c) changing the temperature of the recombinant host cell from: (i) a permissive temperature to a restrictive temperature, or (ii) a restrictive temperature to a permissive temperature, and (d) controlling the reproducibility of the method for regulating the expression of a polypeptide or untranslated RNA molecule.

The present invention also provides methods for producing polypeptides and RNA molecules, the methods comprising introducing one or more nucleic acid molecules or expression systems of the invention into recombinant host cells, expanding the cells at a non-permissive temperature and incubating these cells at a permissive temperature. In a related aspect, the invention provides purified polypeptides and RNA molecules produced according to the methods of the present invention. As used herein, the term "purified" used in reference to a molecule means that the concentration of the molecule being purified has been increased relative to molecules associated with it in its natural environment. Naturally associated molecules include polypeptides, nucleic acids, lipids and sugars but generally do not include water, buffers, and reagents added to maintain the integrity or to facilitate the purification of the molecule. For example, even if mRNA is diluted with an aqueous solvent during oligo dT column chromatography, mRNA molecules are purified by this chromatography if naturally associated nucleic acids and other biological molecules do not bind to the column and are separated from the subject mRNA molecules.

The method of the invention may further comprise recovering the polypeptide or untranslated RNA molecule. Polypeptides produced using the method and the nucleic acid molecules and expression systems of the invention can be recovered and purified from recombinant cell cultures by methods known in the art including ammonium sulfate precipitation, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography, and high performance liquid chromatography. Methods for purifying proteins are described in numerous sources (see, e.g., Celis, J., ed., CELL BIOLOGY, Academic Press, 2nd edition, (1998)).

Untranslated RNA molecules produced using the nucleic acid molecules and expression systems of the invention can be recovered and purified from recombinant cell cultures by methods known in the art (see, e.g., Celis, J., ed., CELL BIOLOGY, Academic Press, 2nd edition, (1998)). Methods for recovering and/or purifying RNA molecules include phenol/chloroform extraction, digestion with DNAs followed by precipitation of the undigested RNA molecules, and column chromatography (e.g., oligo dT column chromatography). Further, RNA molecules can be separated from other cellular material using the single step guanidinium-thiocyanate-phenol-chloroform method described in Chomczynski and Sacchi, Anal. Biochem. 162:156 159 (1987).

Depending on the molecule, which is expressed, it may be obtained either from the culture supernatant or by lysing the recombinant host cells.

It will be readily apparent to one of ordinary skill in the relevant arts that other suitable modifications and adaptations to the methods and applications described herein are obvious and may be made without departing from the scope of the invention or any embodiment thereof. Having now described the present invention in detail, the same will be more clearly understood by reference to the following examples, which are included herewith for purposes of illustration only and are not intended to be limiting of the invention.

EXAMPLES

EXAMPLE 1

pCep derieved VECTORS AND CONSTRUCTS

The vectors of the pCEP system are derived from the commercial available vector pCEP4 (Invitrogen, Cat. No. V004-50). In pCEPpu (SEQ ID NO: 1, Fig. 1) the hygromycin resistance marker in pCEP4 was exchanged by the puromycin resistance marker (Wuttke et al., J Biol Chem, 2001, 39, 36839-48). To generate a version expressing EGFP, pEGFP-N1 (Clontech) was digested with HindIII and NotI and the 779 bp fragment containing the EGFP open reading frame was introduced into the HindIII and NotI sites of pCeppu. The resulting construct was named pCep-EGFP. To generate a

pCep version expressing EPO, the cDNA of EPO was amplified by PCR using the oligonucleotides EPO-ORF-for (SEQ ID NO: 2) and EPO-ORF-rev (SEQ ID NO: 3) from a vector containing the cDNA of EPO . The resulting PCR fragment was digested with HindIII and XhoI and cloned into the HindIII and XhoI sites of pCEPpu leading to pCep-EPO.

Vectors and constructs of the pCytTS-EBV system

Vectors of the pCytTS-EBV system containing either the EBV origin of replication (OriP) alone (pCytTS-OriP vector) or in conjunction with the replication initiation factor EBNA1 (pCytTS-OPE vector), and these vectors with EPO as an insert (pCytTS-EPO-OriP, pCytTS-EPO-OPE) have been described in details in a previous application (PCT/EP03/09291) of the present assignee.

EXAMPLE 2

STABILITY OF EXPRSSION IN DIFFERENT CELL LINES USING THE pCEP VECTOR SYSTEM

To test the stability of episomaly replicating vectors in different human cell lines, pCep-EGFP was introduced into different cell lines of human origin and EGFP expression was monitored over several months. Briefly, 293-EBNA cells (Invitrogen, Cat. No. R620-07), 293F (Invitrogen, Cat. No. 11625) and the serum free adapted 293-H (Invitrogen, Cat. No. 11631) human embryonic kidney cells were transfected with pCep-EGFP using Lipofectamin 2000 (BRL) according to the manufacturer's recommendation. One day after transfection cells were split 1 to 2.5 or 1 to 4 in the presence of puromycin (0.8 µg/ml) After 2-5 days of selection all non transfected cells had died and the resistant semi-stable cell population was further expanded. Resistant cell populations were regularly passaged under selective pressure over several months. At different time points cells were harvested by trypsinisation and the percentage of EGFP positive cells was determined by flow cytometry (Fig.2). In conclusion, the episomaly replicating pCep system showed high stability in different cell lines of human origin (293 EBNA, 293F, 293H), as documented by the high percentage of EGFP positive cells after several weeks of culture.

EXAMPLE 3

STABILITY OF THE PCYTTS-EBV SYSTEM IN IN 293 EBNA CELLS

To test the stability of the temperature inducible, episomaly replicating CytTS-EBV system, pCytTS-EPO-OriP was introduced into 293 EBNA cells and the inducible expression of EPO was monitored during the first 5 weeks after transfection and selection. Briefly, 293 EBNA cells were transfected with pCytTS-EPO-OriP using Lipofectamine 2000 (BRL) according to the manufacturer's recommendation. One day after transfection cells were split 1 to 4 in the presence of 0.8 μg/ml puromycin. After 2-5 days all non transfected cells had died and detached from the plates. The semi-stable cell population was further expanded under selective pressure. To assess inducible EPO expression, resistant cell populations were trypsinized and two T25 flasks were seeded with three million cells each. Once the cells had attached one T25 flask was kept at 37 °C (uninduced) and one was shifted to 29 °C for induction purposes. Cell culture supernatants were collected from both plates seven days after induction and the EPO concentration in the different samples was determined by ELISA (EPO ELISA kit, R&D Systems, Minneapolis, USA). A backup plate of the resistant cell pool was kept in culture and passaged 1 to 10 twice weekly. To determine stability of EPO expression over several weeks, cells from the backup plate were seeded at intervals of one or two weeks in T25 flask as described above. Cells were induced for 7 days and EPO expression was determined by ELISA (Fig. 3). EPO expression increased during the first 3 weeks of the experiment and remained high over the complete experiment. Hence, these results demonstrate that the episomaly replicating pCytTS-EPO-EBV is maintained over several weeks without loss of expression of the gene of interest. Furthermore the CytTS system is still tightly regulated, since hardly any EPO expression could be detected at 37 °C at all time points.

EXAMPLE 4

RAPID EXPRESSION OF RECOMBINANT PROTEINS USNG A SELFREPLICATING SYSTEM IN BATCH FERMENTATION

In a first step 293-EBNA cells were adapted to serum-free growth in 293 SFMII medium, since this allows to speed up to process from transfection to the production of recombinant protein in a fermentation process. Such cells can be transfected and selected in the presence of serum and then transferred directly into serum free medium and expanded without difficulties.

To test the episomal expression system in a fermentation process, pCytTS-EPO-OriP was introduced into 293-EBNA cells (Invitrogen), which constitutively express Epstein-Barr virus replication initiation factor EBNA1 and support OriP mediated DNA replication. Briefly, the day before transfection a 6-well dish was seeded with 0.5 x 10^6 cells/ml. The adherent cells were then transfected with 4 µg DNA of pCytTS-EPO-OriP using Lipofectamine 2000 (BRL) according to the manufacturer's recommendation. 36 h after transfection cells were split into a T-75 flask in presence of 0.8 µg/ml puromycin and 250 µg/ml geneticin. After 5days all non transfected cells had died and selection was completed. The stable cell pool was reseeded in a T75 flask in serum-free medium. Two days later suspension cells were transferred to a baby spinner in a total volume of 40 ml and grown at 37°C at 80 rpm. At a cell density of 1x 10^6 cells/ml cells were seeded in a 1 l disposable bioreactor chamber (CELLBAG CB2L (standard), Cat. No. 502010, Wave Biotech AG) in a total volume of 150 ml at 0.3 x 10^6 cells/ml. Two and five days after inoculation medium was expanded to a final volume of 0,7 l. The cell suspension was then transferred to a 10 l disposable bioreactor chamber (CELLBAG CB20L (standard), Cat. No. 520020, Wave Biotech AG) in a total volume of 1,4 l. Cell suspension was gradually expanded to 7,9 l and the reactor was induced at 29 °C at a cell density of 0.75x10^6 cells/ml. Two days later again 1 l medium was added (total bioreactor volume: 8,9 l). EPO expression and cell number was monitored daily during 5 days after induction (see Fig. 4). After 5 days of induction an EPO expression of 1.5 mg/l was measured, demonstrating, that the episomaly replicating vector system can be used for large scale fermentation.

EXAMPLE 5

RAPID EXPRESSION OF RECOMBINANT PROTEINS USNG A SELFREPLICATING SYSTEM IN PERFUSION FERMENTATION

To test the episomal expression system in a perfusion fermentation process, pCytTS-EPO-OPE was introduced into 293-EBNA cells (Invitrogen), which constitutively express Epstein-Barr virus replication initiation factor EBNA1 and support OriP mediated DNA replication. Briefly, adherent growing 293 EBNA cells in a T75 flask were transfected with 40 µg of pCytTS-EPO-OPE DNA using Lipofectamine 2000 (BRL) according to the manufacturer's recommendation. 36 h after transfection cells were passaged under puromycin selection (0.8 µg/ml puromycin) into a T150 flask. After 5 days of selection all non transfected cells had died. The obtained semi-stable cell pool was cryopreserved. After thawing and initial propagation in T-Flasks with medium supplemented with serum, the stable cell pool was expanded into a spinner flask in a total volume of 200 ml in serum-free medium. At a cell density of 0.44 x 10^6 cells/ml cells were spun down and resuspended in 200 ml of fresh serum-free medium. A 1 l disposable bioreactor chamber (CELLBAG CB2L/PER (perfusion), Cat. No. 502020, Wave Biotech AG) was inoculated with these 200 ml. During 8 days the volume was gradually expanded up to 1 l. At a cell density of 1 x 10^6 cells/ml a perfusion run was started with a two fold medium exchanges every day in order to increase the cell density. After 6 days of perfusion fermentation a cell density of 3.4 x 10^6 cells/ml had been reached and the bioreactor was induced at 29°C. EPO expression and total cell numbers were monitored daily during 4 days (see Fig.4). The maxium EPO expression (5.1 mg EPO/l) was obtained 3 days after induction. This is a level 3.5 fold higher than with the batch method (see Fig.5). Hence, high productivities were achieved in a perfusion fermentation process in less than 3 weeks.

EXAMPLE 6

CONSTRUCTION OF VECTORS OF THE PCEP AND THE PCYTTS SYSTEM CONTAINING THE GLUTAMINE SYNTHETASE AS SELECTABEL MARKER

The glutamine synthetase gene is cloned either from hamster cells (CHO-K1, Chinese hamster ovary, ATCC, Cat. No. CCL-61; BHK21, Hamster Syrian kidney cells, DSMZ, Cat. No. ACC 61) or human cells (HEK 293 cells (ATCC, Cat. No. CRL-1573; Hela ATCC, Cat. No. CCL-2; Raji cells, ATCC, Cat. No. CCL-86; 293 EBNA , Invitrogen, Cat. No. R62007; and 143B cells, ATCC, Cat. No. CRL-8303) by the method of reverse transcription. Total RNA or cytoplasmic RNA is isolated from the cells using the RNeasy Kit (Qiagen, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311) according to the manufactures recommendation. Enrichment of poly(A)+ RNA from total RNA can be done by using the Oligotex mRNA Kit (Qiagen, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311). The resulting RNA can be treated with DNaseI to remove residual traces of DNA. cDNA synthesis is performed in the first step using the ThermoScript RT-PCR System (Invitrogen, Cat. No. 11146-024) with either total RNA or poly(A)+ -selected RNA primed with oligo(dT) or random primers according to the manufactures recommendation. Treatment of cDNA with RNase H to remove the complementary RNA prior to PCR is optional. In the second step, PCR is performed using primers specific for the hamster or human glutamine synthetase. To amplify the cDNA of hamster and human glutamine synthetase the following oligonucleotides are used: hamGSfor (SEQ ID NO: 4), ham GSrev (SEQ ID NO: 5), hGSfor (SEQ ID NO: 6) hGSrev (SEQ ID NO: 7). To introduce the hamster or human glutamine synthetase into pCEP family vectors, the PCR fragments are digested with *Hind*III and *Xba*I in the appropriate buffer and subcloned into the Hind III and XbaI sites of pGL3-Promotervector (Promega, Cat. No. E1761). The resulting vector pGL3-GS is digested with *Sfi*I and *Cla*I to release the glutamine synthetase cassette including the SV40 polyA and parts of the SV40 promoter. The glutamine synthetase cassette is then cloned into pCEPpu via a three fragment ligation. Vector pCEPpu is digested with ClaI and ScaI to release a 2700 bp fragment and with *Sca*I and *Sfi*I to release a 5800 bp fragment. These two fragments are ligated with the

glutamine synthetase fragment resulting in vector pCEPgs, where the puromycin resistance marker is replaced by the glutamine snythetase cassette.

To replace the puromycin resistance marker with the glutamine synthetase cassette in vector pCytTS2.1, the glutamine synthetase cassette is isolated from *Mlu*I and *Bam*HI digested vector pGL3-GS. The fragment is treated with Klenow polymerase in order to get blunt ends and ligated into *Mlu*I digested vector pCytTS2.1 treated with Klenow polymerase.

To combine the resulting pCytTS2.1gs vector with the self replicating EBNA system either the origin of replication alone (pGEMT-OriP) or together with the replication initiation factor (pGemT-OriP) can be excised by *Not*I digestion and transferred into the *Not*I site of pCytTS2.1gs.

Vectors of the present invention containing the glutamine synthetase gene as selection marker are transfected in different cell lines as described above (EXAMPLE 3-5). Mammalian cell lines that do not express GS (glutamine synthetase) cannot survive without added glutamine in the medium. For these cell lines, a transfected GS gene can function as a selectable marker by permitting growth in a glutamine-free medium (e.g. mouse myeloma NS0 cell line). For other cell lines (e.g. CHO-K1) producing endogenous GS, methionine sulphoximine (MSX) has to be added as selective agent to the glutamine-free medium.

EXAMPLE 7

GMP compliant process for the production of recombinant proteins with an episomal expression system

The great advantage of episomal expression systems as described in the examples above, is the rapid generation of a semi-stable cell population which displays high level of protein expression. As opposed to classical expression systems based on stable integration of an expression plasmid into the genome of a host cell, episomal expression systems allow to dramatically reduce the timelines for the large scale production of proteins. This approach is of great value for the production of pharmaceutical proteins under good manufacturing practice (GMP) for the generation of proteins for clinical trials. For these

type of processes, it is crucial to get reproducibility from batch to batch. Reproducibility of the semi-stable, episomal expression system is achieved by constant monitoring of several different parameters over time. Typically the productivity, growth rates and average amounts of episomes per cell are measured at different time points after the generation of a semi-stable cell population. Analysis of these parameters allows for accurate monitoring of a process which is essential for the production of clinical material under GMP conditions.

Briefly, cells are transfected with a plasmid competent for episomal replication encoding a protein to be expressed under GMP conditions like for example a monoclonal antibody. In the case of a monoclonal antibody both the heavy chain and the light chain are expressed under the control of strong mammalian promoters on the same plasmid competent for episomal replication (e.g. a plasmid containing the EBV origin of replication (OriP) and the EBV replication initiation factor EBNA1). Cells are transfected and selected as described in examples 4 and 5. Depending on the size of the fermentation batch and the amounts of desired cell divisions before harvesting of recombinant proteins the amount of initially transfected cells is chosen. At defined time points different parameters are measured. To asses proliferation rates of transfected cells different methods are used depending on the growth conditions of the cells. During the adherent growth phase, defined amounts of cells are seeded and counted at each passage. Based on the amount of seeded cells and the amount of cells at the moment of the next passaging, the growth rates can be determined. Under conditions were cells are grown in suspension, aliquots of cells are taken at different time points and the proliferation rates are determined based on the increase of cell numbers over time.

EXAMPLE 8

Assessment of productivities over time

To asses the productivities over time, defined amounts of cells are harvested at the different time points and resuspended in fresh medium either under conditions of adherent cell growth or under conditions were cells are grown in suspension depending on the growth conditions of the process to be monitored. At defined time points after the inoculation (typically after 5-9 days) of the small cultures for the monitoring of protein

production aliquots of medium are harvested and the content of recombinant protein is determined by quantitative dot blots, quantitative western blots or by sandwich ELISA. For the quantification by quantitative dot blots, defined amounts of supernatants and serial dilutions of a standard of known concentration are blotted onto a nitrocellulose membrane. Then membranes are blocked by the incubation with PBS/0.05 % Tween20 supplemented with 1 % milk powder for 1 hour at room temperature. After extensive washes with PBS/0.05 % Tween20, membranes are incubated with specific antibodies against the produced recombinant protein for 2 hours at room temperature. After extensive washes with PBS/0.05 % Tween20, membranes are stained with an appropriate second stage antibody (e.g. an antibody which recognizes the primary, protein specific antibody) labelled with horse radish peroxidase for 1 hour at room temperature. After extensive washes of the membrane with PBS/0.05 % Tween20 bound antibodies are visualized by enhanced chemiluminescence (ECL) according to manufacturer's recommendations (Amersham). Productivities are determined based on the comparison of the intensity of the supernatants compared to the standard of known concentrations. Alternatively supernatants and standards with known concentrations are separated on SDS page gels. Then proteins are transferred to nitrocellulose membranes by Westernbloting and the membranes are processed as described above for the dot blot. Based on the intensity of the bands, in comparison to a known standard, the amount of recombinant protein in the supernatant can be determined.

Furthermore productivities are monitored by sandwich ELISA (Enzyme Linked Immono Sorbent assay). Briefly Microtiter polystyrene plates (#442404, Nalge-Nunc International) are pre-coated overnight at 4°C with an appropriate, protein specific antibody at a concentration of 1-10 µg/ml in 0.1M sodium carbonate, pH 9.6. The plates are then washed and incubated with PBS/0.05% Tween20 supplemented with 1% milk powder for 1 h at room temperature for blocking purposes. Then appropriate dilutions of the samples are made, added to the plates and incubated for 1 hour at room temperature. The plates are then washed several times with washing buffer in order to remove unbound proteins. The plates are then stained with a second protein specific antibody for 1 hour at room temperature. After extensive washing with PBS/0.05% Tween20, plates are incubated with an appropriate second stage antibody (specific antibody which recognises the second antibody) for 45 minutes at room temperature. After extensive washing, the bound complexes are visualized using the chromogenic substrate tetramethylbenzidine

(TMB, #T0440, Sigma-Aldrich Chemie GmbH). Concentrations of recombinant proteins are estimated based on a standard with known concentration.

EXAMPLE 9

Quantification of episomes

In order to quantify the amounts of episomes in the semi-stable cell populations over time different approaches are used. To determine average copy number in the cell population, either southern blot procedures or quantitative PCR as described below is used. To monitor the maintenance and copy number of the episome by Southern blot a procedure described previously is used (Wu et al.; Journal of virology, 76, 5, 2480-2490). Briefly, $5x10^6$ cells from each culture are harvested at each time point and lysed by the method of Hirt (Hirt, B., J. Mol. Biol. 1967, 26:365–369). Low-molecular weight DNA is then isolated as described by Ceccarelli and Frappier (Ceccarelli, D. F. J., and L. Frappier, J. Virol. 2000, 74:4939–4948), digested with XhoI and DpnI, separated by agarose gel electrophoresis, Southern blotted, and probed with ^{32}P-labeled probes specific for the vectors. Linearized plasmid bands are visualized by autoradiography and quantified by phosphor-imager analysis using ImageQuant software (Molecular Dynamics) and compared to a standard of known copy number. In an other approach the average number of episome in the cell population is determined by quantitative PCR using specific primers for the episome. Briefly, cellular DNA (Genomic and episomal DNA) is purified from transfected cells using the Tri Reagent method (US patents 4,843,155 and 5,346,994 and international patents). From a cell suspension, $2x10^6$ cells were pelleted by centrifugation, resuspended in phosphate-buffered saline (PBS) and pelleted again. The cell pellet is lysed in 500 µl Tri Reagent® (Sigma, Saint Louis, MO; Molecular Research Center, Inc. Cincinnati, OH). The large DNA molecules are sheared by intensive vortexing for one minute. After incubation at room temperature for 5 min., 100 µl of chloroform or 50 µl of bromocholopropane (BCP) are added and the samples are vortexed intensively, incubated for 10 minutes at room temperature and then centrifuged at 16,000 x g at 4°C for 15 min. The aqueous top phase containing RNA but no DNA and protein is aspirated and discarded. 150 µl of ethanol is added to the combined lower fractions containing DNA and protein. Upon mixing and incubation at room temperature for 2

min., the precipitated DNA is pelleted by centrifugation with 2,000 x g at 4°C for 5 min. The supernatants is removed and the DNA precipitate is soaked in 500 µl sodium citrate solution (0.1 M sodium citrate in 10% ethanol) for 30 min. Centrifugation at 2,000 x g and soaking in sodium citrate solution is repeated once. Subsequently, the DNA precipitate is soaked in 500 µl of 75% ethanol at room temperature for at least 20 min. After centrifugation at 2,000 x g at 4°C for 5 min., all liquid is removed and the DNA precipitate is dried. The dried DNA is dissolved in 100 µl of 8 mM NaOH in H_2O and subsequently neutralized with 6.6 µl of 0.1M N-2-Hydroxyethylpiperazine-N'-2-ethanesulfonic acid (HEPES) in H_2O. DNA concentration is measured by absorbance at 260 nm of an 1/25 dilution of the sample. The copy number of plasmids per cell is determined by quantitative real-time PCR with specific primers for the episomaly replicating plasmid. More specifically, primers which amplify a part of the origin of replication (OriP) are chosen. Briefly, PCR reaction contained 10 ng BamH I digested DNA, forward (SEQ ID NO: 9) and reverse primer (SEQ ID NO: 10) at 0.3 µM each, KCl (50 mM), Tris-HCl pH 8.4 (20 mM), $MgCl_2$ (3 mM), formamide (3%), dATP, dCTP, dGTP, dTTP (0.2 mM each), Platinum® Taq DNA polymerase (0.625 units; Invitrogen), Fluorescein (5 nM), and Sybr® Green I (0.1 x; Molecular Probes, Eugene, OR). Polymerase chain reaction is performed in an iCycler instrument (Bio-Rad, Hercules, CA). After an initial incubation denaturation step of 90 seconds at 95°C temperature cycling is performed between 94°C for 20 seconds and 58°C for 40 seconds. In total 40 such cycles are performed. Fluorescence is measured at each 58°C step. For quantification purposes, a dilution series of known numbers of DNA molecules is amplified at the same time using the same primer pair. The number of plasmids is calculated with the iCycler™ iQ Optical Sytem Software, version 3.0a and the number of plasmids per cell is calculated.

Having now fully described the present invention in some detail by way of illustration and example for purposes of clarity of understanding, it will be obvious to one of ordinary skill in the art that the same can be performed by modifying or changing the invention within a wide and equivalent range of conditions, formulations and other parameters without affecting the scope of the invention or any specific embodiment thereof, and that such modifications or changes are intended to be encompassed within the scope of the appended claims.

All publications, patents and patent applications mentioned in this specification are indicative of the level of skill of those skilled in the art to which this invention pertains, and are herein incorporated by reference to the same extent as if each individual publication, patent or patent application was specifically and individually indicated to be incorporated by reference.

CLAIMS:

1. A method for producing a polypeptide or untranslated RNA molecule, said method comprising the steps of

(A) providing at least one host cell;

(B) introducing at least one nucleic acid molecule into said at least one host cell to produce at least one recombinant host cell, said nucleic acid molecule comprising

(a) a first polynucleotide element capable of replicating said at least one nucleic acid molecule in said at least one host cell; and

(b) at least one second polynucleotide element selected from the group consisting of

i. an open reading frame encoding a polypeptide of interest;

ii. a nucleotide sequence complementary to all or a part of the open reading frame of (i); and

iii. a nucleotide sequence encoding an untranslated RNA molecule or complement thereof;

(C) selecting at least one stable recombinant host cell from said at least one host cell;

(D) culturing said at least one recombinant host cell under conditions suitable for expression of said polypeptide or untranslated RNA molecule; and

(E) controlling the reproducibility of said method for producing said polypeptide or said untranslated RNA molecule as a function of time.

2. The method of claim 1, wherein said steps are performed consecutively.

3. The method of any of the preceding claims, wherein said controlling the reproducibility of said method is effected by a way selected from

(a) measuring the amount of said polypeptide or untranslated RNA molecule per cell expressed by said at least one recombinant host cell as a function of time; and

(b) determining the number of said at least one nucleic acid molecule being in said at least one recombinant host cell as a function of time.

4. The method of any of the preceding claims, wherein said controlling the reproducibility of said method is effected by way of measuring the amount of said polypeptide or untranslated RNA molecule per cell expressed by said at least one recombinant host cell as a function of time.

5. The method of claim any one of the preceding claims, wherein said controlling the reproducibility of said method is effected by way of determining the number of said at least one nucleic acid molecule being in said at least one recombinant host cell as a function of time.

6. The method of claim 5 wherein said determining the number of said nucleic acid molecule being in said recombinant host cell is performed by quantitative PCR.

7. The method of any one of the preceding claims further comprising the step of adapting said at least one host cell to serum-free growth.

8. The method of claim 7 wherein said step of adapting said at least one host cell to serum-free growth is effected prior to said introducing said nucleic acid molecule into said host cell to produce a recombinant host cell.

9. The method of any one of the preceding claims further comprising recovering said polynucleotide produced by said recombinant host cell.

10. The method of any one of the preceding claims wherein said host cell is a eukaryotic host cell.

11. The method of any one of the preceding claims wherein said nucleic acid molecule is an episomally replicating expression vector.

12. The method of any one of the preceding claims wherein said first polynucleotide element comprises

(a) a first polynucleotide sequence comprising an origin of replication; and

(b) a second polynucleotide sequence encoding a replication initiation factor capable of recognizing said origin of replication.

13. The method of claim 12, wherein said origin of replication is derived from a virus that allows for episomal replication, preferably form a Herpesvirus or Papovavirus.

14. The method of claim 13, wherein said origin of replication is oriP derived from Epstein-Barr virus (EBV).

15. The method of any one of claim 12 to 14 , wherein said replication initiation factor is derived from a virus that allows for episomal replication, preferably form a Herpesvirus or Papovavirus.

16. The method of claim 15, wherein said replication initiation factor is EBNA-1 derived from Epstein-Barr virus (EBV).

17. The method of any one of claims 12 to 16, wherein said replication initiation factor is capable of operating as a plasmid maintenance factor.

18. The method of any one of claims 12 to 17, wherein said replication initiation factor is constitutively expressed in said at least one host cell.

19. The method of any one of claims 12 to 17, wherein said replication initiation factor is inducibly expressed in said at least one host cell.

20. The method of any one of claims 12 to 19, wherein said replication initiation factor and said origin of replication are derived from different organisms or viruses.

21. The method of any one of the preceding claims wherein said nucleic acid molecule further comprises a selection marker, and wherein preferably said selection marker confers resistance to puromycin or wherein said selection marker is glutamine synthetase.

22. The method of any one of the preceding claims wherein said selecting at least one stable recombinant host cell is performed by growing cells in the presence of puromycin or in the absence of glutamine.

23. The method of any one of the preceding claims, wherein said second polynucleotide element encodes the light chain and/or heavy chain of an antibody, and antigene, a cytokine, an enzyme, a lymphokine, a tumor necrosis factor, an interferon, a toxic polypeptide, a prodrug or a converting enzyme.

24. The method of any one of the preceding claims, wherein said introducing a nucleic acid molecule into said host cell to produce at least one recombinant host cell is performed by transfection with said nucleic acid molecule.

25. The method of any one of the preceding claims, wherein said culturing said recombinant host cell under conditions suitable for expression of said polypeptide or untranslated RNA molecule is carried out under serum-free conditions.

26. The method of any one of the preceding claims, wherein said culturing said recombinant host cell under conditions suitable for expression of said polypeptide or untranslated RNA molecule is performed in a reactor.

27. A method for producing a polypeptide or untranslated RNA molecule conforming to Good Manufacturing Practice (GMP), comprising carrying out the steps according to any one of the preceding claims.

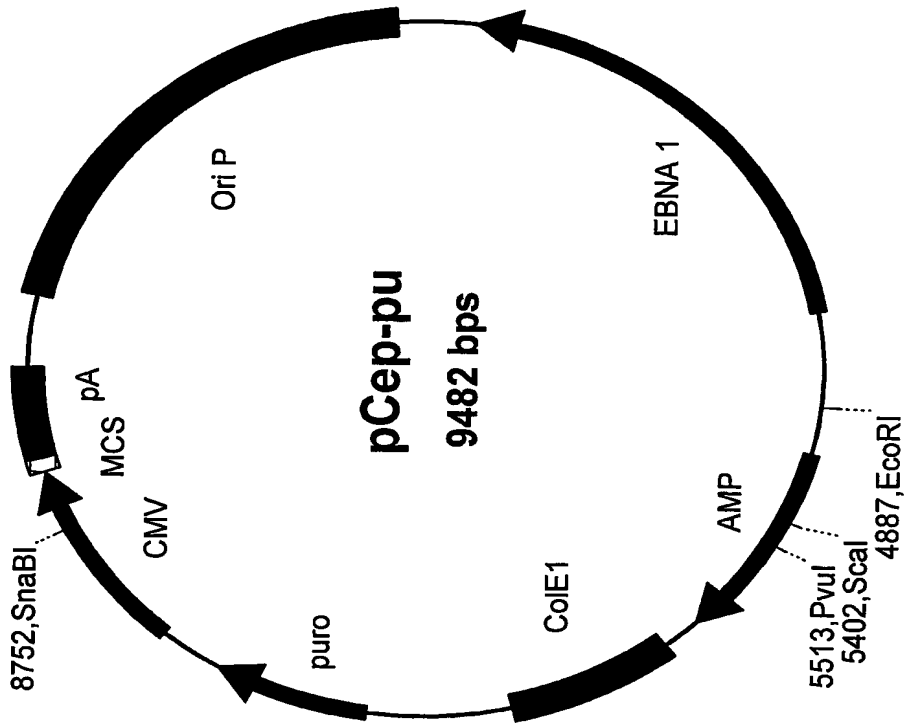
pCep-pu
9482 bps
Ori P
EBNA 1
pA
MCS
CMV
AMP
puro
ColE1
8752,SnaBI
5513,PvuI
5402,ScaI
4887,EcoRI

Fig. 1

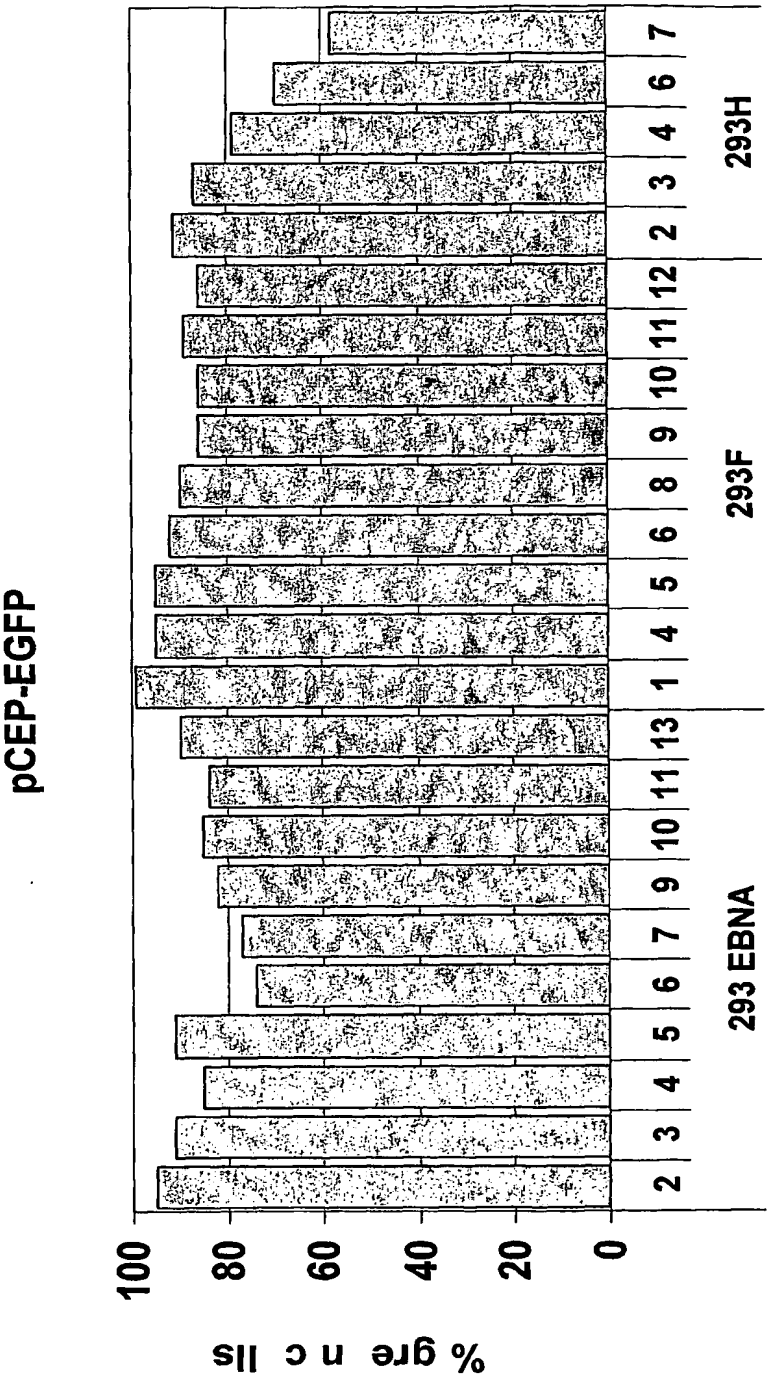
pCEP-EGFP
% gre n c lls
100
80
60
40
20
0
2 3 4 5 6 7 9 10 11 13 1 4 5 6 8 9 10 11 12 2 3 4 6 7
293 EBNA
293F
293H
weeks in culture

Fig. 2

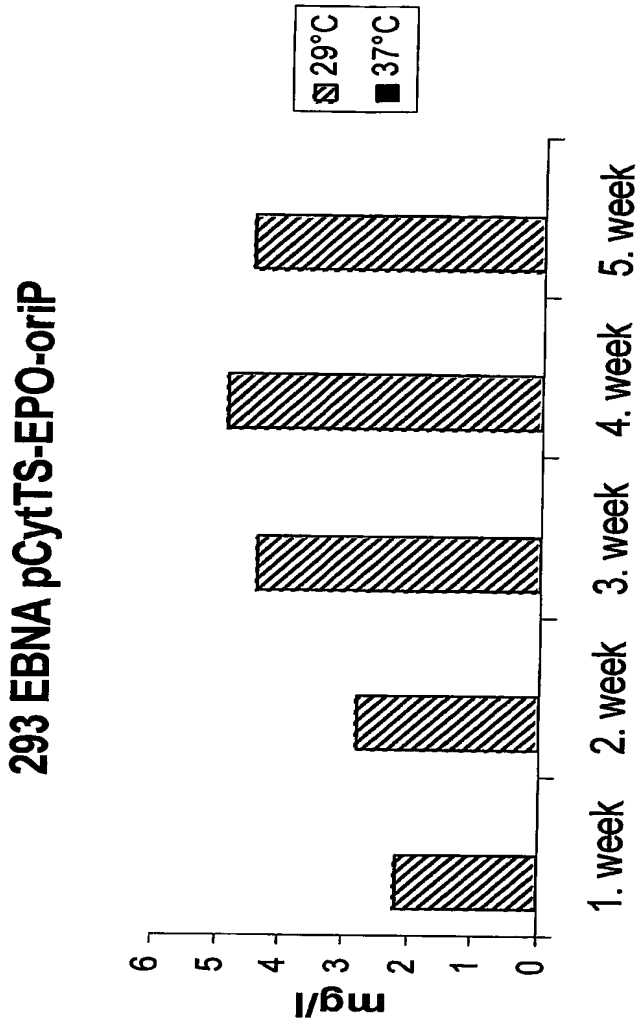
293 EBNA pCytTS-EPO-oriP
29°C
37°C
mg/l
0
1
2
3
4
5
6
1. week
2. week
3. week
4. week
5. week

Fig. 3

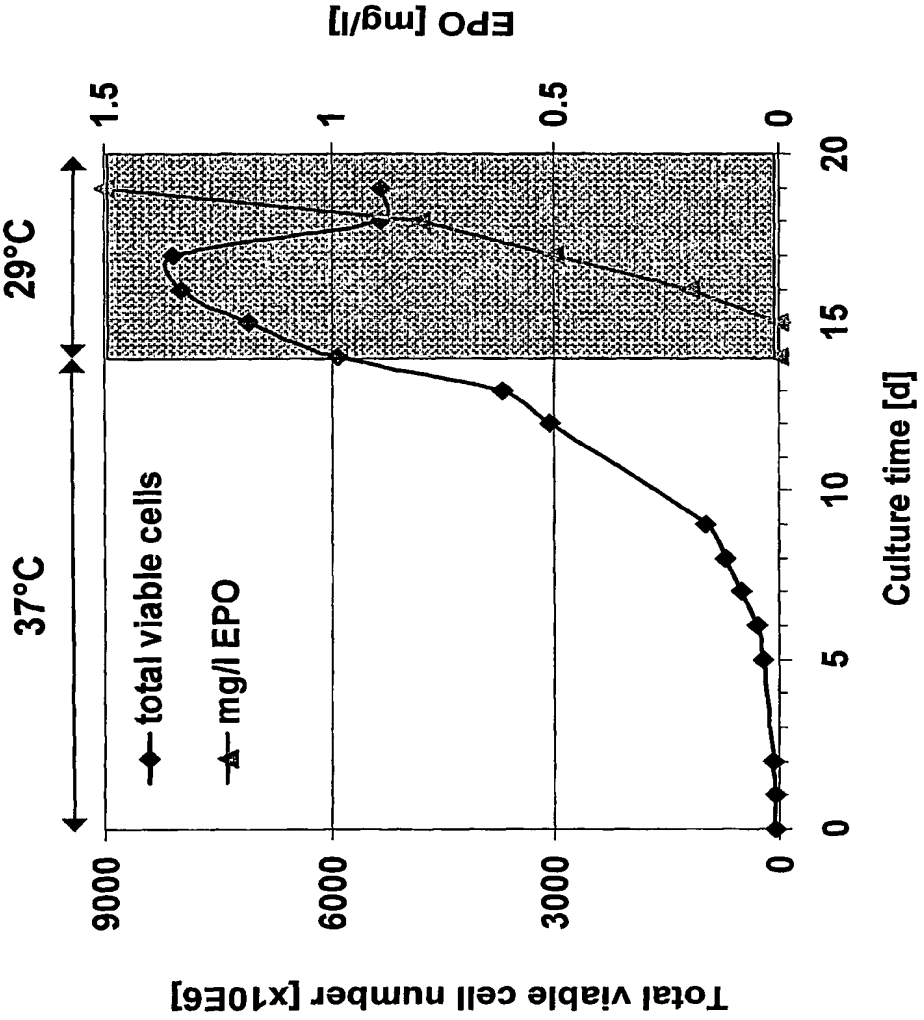
EPO [mg/l]
1.5
1
0.5
0
29°C
37°C
total viable cells
mg/l EPO
9000
6000
3000
0
0
5
10
15
20
Culture time [d]
Total viable cell number [x10E6]

Fig. 4

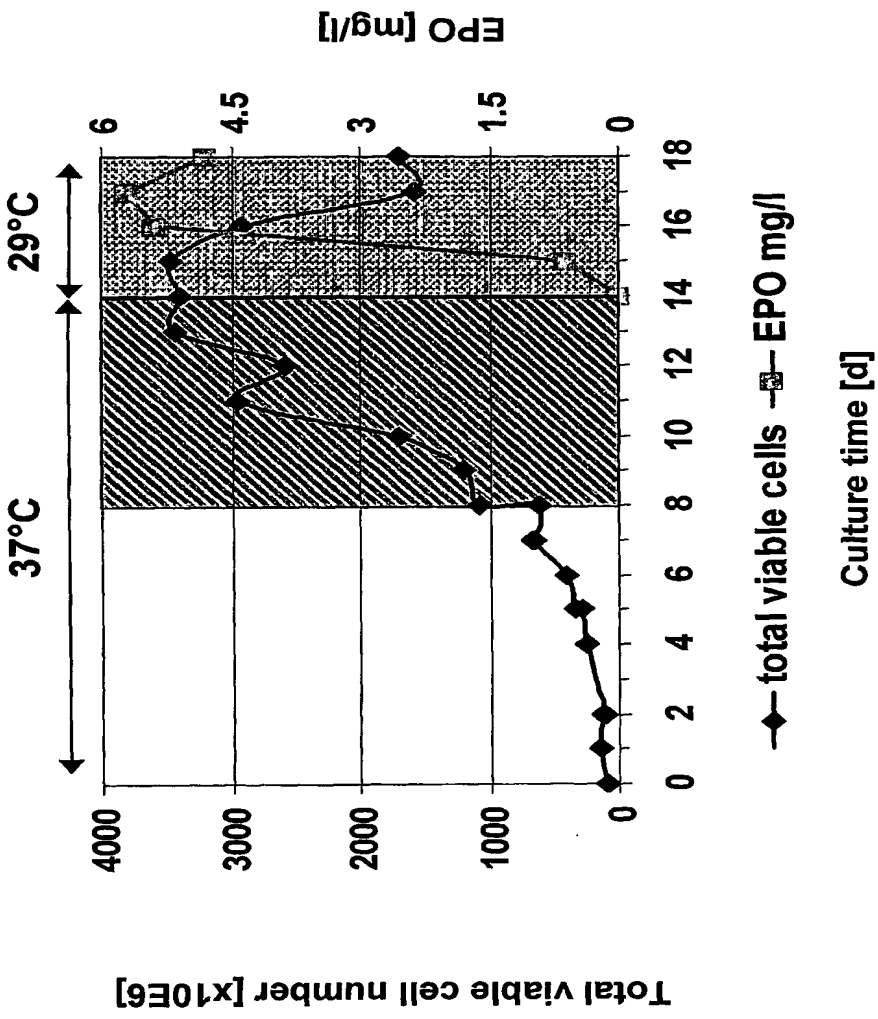
EPO [mg/l]
6
4.5
3
1.5
0
29°C
37°C
4000
3000
2000
1000
0
0 2 4 6 8 10 12 14 16 18
total viable cells
EPO mg/l
Culture time [d]
Total viable cell number [x10E6]

Fig. 5

SEQUENCE LISTING

<110> Cytos Biotechnology AG

<120> Method for Protein Production

<130> PA037WO

<150> US 60/432,245
<151> 2002-12-11

<160> 10

<170> PatentIn version 3.2

<210> 1
<211> 9482
<212> DNA
<213> Artificial Sequence

<220>
<223> cDNA

<400> 1
ggatcgatcc ccgccgccgg acgaactaaa cctgactacg gcatctctgc cccttcttcg 60
cggggcagtg catgtaatcc cttcagttgg ttggtacaac ttgccaactg ggccctgttc 120
cacatgtgac acgggggggg accaaacaca aaggggttct ctgactgtag ttgacatcct 180
tataaatgga tgtgcacatt tgccaacact gagtggcttt catcctggag cagactttgc 240
agtctgtgga ctgcaacaca acattgcctt tatgtgtaac tcttggctga agctcttaca 300
ccaatgctgg gggacatgta cctcccaggg gcccaggaag actacgggag gctacaccaa 360
cgtcaatcag aggggcctgt gtagctaccg ataagcggac cctcaagagg gcattagcaa 420
tagtgtttat aaggcccccc tgttaaccct aaacgggtag catatgcttc ccgggtagta 480
gtatatacta tccagactaa ccctaattca atagcatatg ttacccaacg ggaagcatat 540
gctatcgaat tagggttagt aaaagggtcc taaggaacag cgatatctcc caccccatga 600
gctgtcacgg ttttatttac atggggtcag gattccacga gggtagtgaa ccattttagt 660
cacaagggca gtggctgaag atcaaggagc gggcagtgaa ctctcctgaa tcttcgcctg 720
cttcttcatt ctccttcgtt tagctaatag aataactgct gagttgtgaa cagtaaggtg 780
tatgtgaggt gctcgaaaac aaggtttcag gtgacgcccc cagaataaaa tttggacggg 840
gggttcagtg gtggcattgt gctatgacac caatataacc ctcacaaacc ccttgggcaa 900
taaatactag tgtaggaatg aaacattctg aatatcttta acaatagaaa tccatggggt 960
ggggacaagc cgtaaagact ggatgtccat ctcacacgaa tttatggcta tgggcaacac 1020
ataatcctag tgcaatatga tactggggtt attaagatgt gtcccaggca gggaccaaga 1080
caggtgaacc atgttgttac actctatttg taacaagggg aaagagagtg gacgccgaca 1140
gcagcggact ccactggttg tctctaacac ccccgaaaat taaacggggc tccacgccaa 1200
tggggcccat aaacaaagac aagtggccac tctttttttt gaaattgtgg agtgggggca 1260
cgcgtcagcc cccacacgcc gccctgcggt tttggactgt aaaataaggg tgtaataact 1320
tggctgattg taaccccgct aaccactgcg gtcaaaccac ttgcccacaa aaccactaat 1380
ggcaccccgg ggaatacctg cataagtagg tgggcgggcc aagatagggg cgcgattgct 1440
gcgatctgga ggacaaatta cacacacttg cgcctgagcg ccaagcacag ggttgttggt 1500
cctcatattc acgaggtcgc tgagagcacg gtgggctaat gttgccatgg gtagcatata 1560

ctacccaaat atctggatag catatgctat cctaatctat atctgggtag cataggctat 1620
cctaatctat atctgggtag catatgctat cctaatctat atctgggtag tatatgctat 1680
cctaatttat atctgggtag cataggctat cctaatctat atctgggtag catatgctat 1740
cctaatctat atctgggtag tatatgctat cctaatctgt atccgggtag catatgctat 1800
cctaatagag attagggtag tatatgctat cctaatttat atctgggtag catatactac 1860
ccaaatatct ggatagcata tgctatccta atctatatct gggtagcata tgctatccta 1920
atctatatct gggtagcata ggctatccta atctatatct gggtagcata tgctatccta 1980
atctatatct gggtagtata tgctatccta atttatatct gggtagcata ggctatccta 2040
atctatatct gggtagcata tgctatccta atctatatct gggtagtata tgctatccta 2100
atctgtatcc gggtagcata tgctatcctc atgcatatac agtcagcata tgatacccag 2160
tagtagagtg ggagtgctat cctttgcata tgccgccacc tcccaagggg gcgtgaattt 2220
tcgctgcttg tccttttcct gcatgctggt tgctcccatt cttaggtgaa tttaaggagg 2280
ccaggctaaa gccgtcgcat gtctgattgc tcaccaggta aatgtcgcta atgttttcca 2340
acgcgagaag gtgttgagcg cggagctgag tgacgtgaca acatgggtat gcccaattgc 2400
cccatgttgg gaggacgaaa atggtgacaa gacagatggc cagaaataca ccaacagcac 2460
gcatgatgtc tactggggat ttattcttta gtgcgggga atacacggct tttaatacga 2520
ttgagggcgt ctcctaacaa gttacatcac tcctgccctt cctcaccctc atctccatca 2580
cctccttcat ctccgtcatc tccgtcatca ccctccgcgg cagccccttc caccataggt 2640
ggaaaccagg gaggcaaatc tactccatcg tcaaagctgc acacagtcac cctgatattg 2700
caggtaggag cgggctttgt cataacaagg tccttaatcg catccttcaa aacctcagca 2760
aatatatgag tttgtaaaaa gaccatgaaa taacagacaa tggactccct tagcgggcca 2820
ggttgtgggc cgggtccagg ggccattcca aaggggagac gactcaatgg tgtaagacga 2880
cattgtggaa tagcaagggc agttcctcgc cttaggttgt aaagggaggt cttactacct 2940
ccatatacga acacaccggc gacccaagtt ccttcgtcgg tagtcctttc tacgtgactc 3000
ctagccagga gagctcttaa accttctgca atgttctcaa atttcgggtt ggaacctcct 3060
tgaccacgat gctttccaaa ccaccctcct tttttgcgcc tgcctccatc accctgaccc 3120
cggggtccag tgcttgggcc ttctcctggg tcatctgcgg ggccctgctc tatcgctccc 3180
gggggcacgt caggctcacc atctgggcca ccttcttggt ggtattcaaa ataatcggct 3240
tcccctacag ggtggaaaaa tggccttcta cctggagggg gcctgcgcgg tggagacccg 3300
gatgatgatg actgactact gggactcctg ggcctctttt ctccacgtcc acgacctctc 3360
cccctggctc tttcacgact tccccccctg gctctttcac gtcctctacc ccggcggcct 3420
ccactacctc ctcgaccccg gcctccacta cctcctcgac cccggcctcc actgcctcct 3480
cgaccccggc ctccacctcc tgctcctgcc cctcctgctc ctgcccctcc tcctgctcct 3540
gcccctcctg cccctcctgc tcctgcccct cctgccccctc ctgctcctgc ccctcctgcc 3600
cctcctgctc ctgccccctcc tgccccctcct cctgctcctg cccctcctgc ccctcctcct 3660
gctcctgccc ctcctgcccc tcctgctcct gcccctcctg cccctcctgc tcctgcccct 3720
cctgccccctc ctgctcctgc ccctcctgct cctgccccctc ctgctcctgc ccctcctgct 3780
cctgccccctc ctgccccctcc tgccccctcct cctgctcctg cccctcctgc tcctgcccct 3840

cctgcccctc ctgcccctcc tgctcctgcc cctcctcctg ctcctgcccc tcctgcccct 3900
cctgcccctc ctcctgctcc tgcccctcct gcccctcctc ctgctcctgc ccctcctcct 3960
gctcctgccc ctcctgcccc tcctgcccct cctcctgctc ctgcccctcc tgcccctcct 4020
cctgctcctg cccctcctcc tgctcctgcc cctcctgccc ctcctgcccc tcctcctgct 4080
cctgcccctc ctcctgctcc tgcccctcct gcccctcctg cccctcctgc ccctcctcct 4140
gctcctgccc ctcctcctgc tcctgcccct cctgctcctg cccctcccgc tcctgctcct 4200
gctcctgttc caccgtgggt ccctttgcag ccaatgcaac ttggacgttt ttggggtctc 4260
cggacaccat ctctatgtct tggccctgat cctgagccgc ccggggctcc tggtcttccg 4320
cctcctcgtc ctcgtcctct tccccgtcct cgtccatggt tatcaccccc tcttctttga 4380
ggtccactgc cgccggagcc ttctggtcca gatgtgtctc ccttctctcc taggccattt 4440
ccaggtcctg tacctggccc ctcgtcagac atgattcaca ctaaaagaga tcaatagaca 4500
tctttattag acgacgctca gtgaatacag ggagtgcaga ctcctgcccc ctccaacagc 4560
ccccccaccc tcatcccctt catggtcgct gtcagacaga tccaggtctg aaaattcccc 4620
atcctccgaa ccatcctcgt cctcatcacc aattactcgc agcccggaaa actcccgctg 4680
aacatcctca agatttgcgt cctgagcctc aagccaggcc tcaaattcct cgtccccctt 4740
tttgctggac ggtagggatg gggattctcg ggacccctcc tcttcctctt caaggtcacc 4800
agacagagat gctactgggg caacggaaga aaagctgggt gcggcctgtg aggatcagct 4860
tatcgatgat aagctgtcaa acatgagaat tcttgaagac gaaagggcct cgtgatacgc 4920
ctatttttat aggttaatgt catgataata atggtttctt agacgtcagg tggcactttt 4980
cggggaaatg tgcgcggaac ccctatttgt ttatttttct aaatacattc aaatatgtat 5040
ccgctcatga gacaataacc ctgataaatg cttcaataat attgaaaaag gaagagtatg 5100
agtattcaac atttccgtgt cgcccttatt cccttttttg cggcattttg ccttcctgtt 5160
tttgctcacc cagaaacgct ggtgaaagta aaagatgctg aagatcagtt gggtgcacga 5220
gtgggttaca tcgaactgga tctcaacagc ggtaagatcc ttgagagttt tcgccccgaa 5280
gaacgttttc caatgatgag cacttttaaa gttctgctat gtggcgcggt attatcccgt 5340
gttgacgccg ggcaagagca actcggtcgc cgcatacact attctcagaa tgacttggtt 5400
gagtactcac cagtcacaga aaagcatctt acggatggca tgacagtaag agaattatgc 5460
agtgctgcca taaccatgag tgataacact gcggccaact tacttctgac aacgatcgga 5520
ggaccgaagg agctaaccgc ttttttgcac aacatggggg atcatgtaac tcgccttgat 5580
cgttgggaac cggagctgaa tgaagccata ccaaacgacg agcgtgacac cacgatgcct 5640
gcagcaatgg caacaacgtt gcgcaaacta ttaactggcg aactacttac tctagcttcc 5700
cggcaacaat taatagactg gatggaggcg gataaagttg caggaccact tctgcgctcg 5760
gcccttccgg ctggctggtt tattgctgat aaatctggag ccggtgagcg tgggtctcgc 5820
ggtatcattg cagcactggg gccagatggt aagccctccc gtatcgtagt tatctacacg 5880
acggggagtc aggcaactat ggatgaacga aatagacaga tcgctgagat aggtgcctca 5940
ctgattaagc attggtaact gtcagaccaa gtttactcat atatacttta gattgattta 6000
aaacttcatt tttaatttaa aaggatctag gtgaagatcc tttttgataa tctcatgacc 6060
aaaatccctt aacgtgagtt ttcgttccac tgagcgtcag accccgtaga aaagatcaaa 6120

ggatcttctt gagatccttt ttttctgcgc gtaatctgct gcttgcaaac aaaaaaacca 6180
ccgctaccag cggtggtttg tttgccggat caagagctac caactctttt tccgaaggta 6240
actggcttca gcagagcgca gataccaaat actgtccttc tagtgtagcc gtagttaggc 6300
caccacttca agaactctgt agcaccgcct acatacctcg ctctgctaat cctgttacca 6360
gtggctgctg ccagtggcga taagtcgtgt cttaccgggt tggactcaag acgatagtta 6420
ccggataagg cgcagcggtc gggctgaacg gggggttcgt gcacacagcc cagcttggag 6480
cgaacgacct acaccgaact gagataccta cagcgtgagc tatgagaaag cgccacgctt 6540
cccgaaggga gaaaggcgga caggtatccg gtaagcggca gggtcggaac aggagagcgc 6600
acgagggagc ttccaggggg aaacgcctgg tatctttata gtcctgtcgg gtttcgccac 6660
ctctgacttg agcgtcgatt tttgtgatgc tcgtcagggg ggcggagcct atggaaaaac 6720
gccagcaacg cggccttttt acggttcctg gcctttgct gcgccgcgtg cggctgctgg 6780
agatggcgga cgcgatggat atgttctgcc aagggttggt ttgcgcattc acagttctcc 6840
gcaagaattg attggctcca attcttggag tggtgaatcc gttagcgagg ccatccagcc 6900
tcgcgtcgaa ctagatgatc cgctgtggaa tgtgtgtcag ttagggtgtg gaaagtcccc 6960
aggctcccca gcaggcagaa gtatgcaaag catgcatctc aattagtcag caaccaggtg 7020
tggaaagtcc ccaggctccc cagcaggcag aagtatgcaa agcatgcatc tcaattagtc 7080
agcaaccata gtcccgcccc taactccgcc catcccgccc ctaactccgc ccagttccgc 7140
ccattctccg ccccatggct gactaatttt ttttatttat gcagaggccg aggccgcctc 7200
ggcctctgag ctattccaga agtagtgagg aggcttttt ggagggtgac cgccacgagg 7260
tgccgccacc atcccctgac ccacgcccct gacccctcac aaggagacga ccttccatga 7320
ccgagtacaa gcccacggtg cgcctcgcca cccgcgacga cgtcccccgg gccgtacgca 7380
ccctcgccgc cgcgttcgcc gactaccccg ccacgcgcca caccgtcgac cccgaccgcc 7440
acatcgaacg cgtcaccgag ctgcaagaac tcttcctcac gcgcgtcggg ctcgacatcg 7500
gcaaggtgtg ggtcgcggac gacggcgccg cggtggcggt ctggaccacg ccggagagcg 7560
tcgaagcggg ggcggtgttc gccgagatcg gcccgcgcat ggccgagttg agcggttccc 7620
ggctggccgc gcagcaacag atggaaggcc tcctggcgcc gcaccggccc aaggagcccg 7680
cgtggttcct ggccaccgtc ggcgtctcgc ccgaccacca gggcaagggt ctgggcagcg 7740
ccgtcgtgct ccccggagtg gaggcggccg agcgcgccgg ggtgcccgcc ttcctggaga 7800
cctccgcgcc ccgcaacctc cccttctacg agcggctcgg cttcaccgtc accgccgacg 7860
tcgagtgccc gaaggaccgc gcgacctggt gcatgacccg caagcccggt gcctgacgcc 7920
cgccccacga cccgcagcgc ccgaccgaaa ggagcgcacg acccggtccg acggcggccc 7980
acgggtccca ggggggtcga cctcgaaact tgtttattgc agcttataat ggttacaaat 8040
aaagcaatag catcacaaat ttcacaaata aagcattttt ttcactgcat tctagttgtg 8100
gtttgtccaa actcatcaat gtatcttatc atgtctggat cgatccgaac cccttcctcg 8160
accaattctc atgtttgaca gcttatcatc gcagatccgg gcaacgttgt tgcattgctg 8220
caggcgcaga actggtaggt atggaagatc tatacattga atcaatattg gcaattagcc 8280
atattagtca ttggttatat agcataaatc aatattggct attggccatt gcatacgttg 8340
tatctatatc ataatatgta catttatatt ggctcatgtc caatatgacc gccatgttga 8400

cattgattat tgactagtta ttaatagtaa tcaattacgg ggtcattagt tcatagccca 8460

tatatggagt tccgcgttac ataacttacg gtaaatggcc cgcctggctg accgcccaac 8520

gacccccgcc cattgacgtc aataatgacg tatgttccca tagtaacgcc aatagggact 8580

ttccattgac gtcaatgggt ggagtattta cggtaaactg cccacttggc agtacatcaa 8640

gtgtatcata tgccaagtcc gcccccctatt gacgtcaatg acggtaaatg gcccgcctgg 8700

cattatgccc agtacatgac cttacgggac tttcctactt ggcagtacat ctacgtatta 8760

gtcatcgcta ttaccatggt gatgcggttt tggcagtaca ccaatgggcg tggatagcgg 8820

tttgactcac ggggatttcc aagtctccac cccattgacg tcaatgggag tttgttttgg 8880

caccaaaatc aacgggactt tccaaaatgt cgtaataacc ccgccccgtt gacgcaaatg 8940

ggcggtaggc gtgtacggtg ggaggtctat ataagcagag ctcgtttagt gaaccgtcag 9000

atctctagaa gctgggtacc ttaaggcgcc agctgatcaa gcttgctagc ggccgctcga 9060

ggccggcaag gccggatcca gacatgataa gatacattga tgagtttgga caaaccacaa 9120

ctagaatgca gtgaaaaaaa tgctttattt gtgaaatttg tgatgctatt gctttatttg 9180

taaccattat aagctgcaat aaacaagtta acaacaacaa ttgcattcat tttatgtttc 9240

aggttcaggg ggaggtgggg aggtttttta aagcaagtaa aacctctaca aatgtggtat 9300

ggctgattat gatccggctg cctcgcgcgt ttcggtgatg acggtgaaaa cctctgacac 9360

atgcagctcc cggagacggt cacagcttgt ctgtaagcgg atgccgggag cagacaagcc 9420

cgtcagggcg cgtcagcggg tgttggcggg tgtcggggcg cagccatgag gtcgactcta 9480

ga 9482

<210> 2
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 2
cttcgacaag cttggcgcgg agatgggggt gcacgaatg 39

<210> 3
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 3
cttcgacctc gagtcatctg tcccctgtcc tgcaggcc 38

<210> 4
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 4
gtcatgaagc ttgccaccgc tcagagcacc ttc 33

<210> 5

<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 5
gacttctaga ctggggcggg gtgggatgaa c 31

<210> 6
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 6
gtcatgaagc ttctcggcga ccagaacacc ttc 33

<210> 7
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 7
gacttctaga ctggggcggg gtgggatgaa c 31

<210> 8
<211> 2200
<212> DNA
<213> homo sapiens

<400> 8
aattcagcaa ggtcgccacg cacaagatca atattaacaa tcagtcatct ctctttagca 60
ataaaaaggt gaaaaattac attttaaaaa tgacaccata gacgatgtat gaaaataatc 120
tacttggaaa taaatctagg caaagaagtg caagactgtt acccagaaaa cttacaaatt 180
gtaaatgaga ggttagtgaa gatttaaatg aatgaagatc taaataaact tataaattgt 240
gagagaaatt aatgaatgtc taagttaatg cagaaacgga gagacatact atattcatga 300
actaaaagac ttaatattgt gaaggtatac tttcttttca cataaatttg tagtcaatat 360
gttcacccca aaaaagctgt ttgttaactt gtcaacctca tttcaaaatg tatatagaaa 420
gcccaaagac aataacaaaa atattcttgt agaacaaaat gggaaagaat gttccactaa 480
atatcaagat ttagagcaaa gcatgagatg tgtggggata gacagtgagg ctgataaaat 540
agagtagagc tcagaaacag acccattgat atatgtaagt gacctatgaa aaaaatatgg 600
cattttacaa tgggaaaatg atgatctttt tctttttttag aaaaacaggg aaatatattt 660
atatgtaaaa aataaaaggg aacccatatg tcataccata cacacaaaaa aattccagtg 720
aattataagt ctaaatggag aaggcaaaac tttaaatctt ttagaaaata atatagaagc 780
atgccatcat gacttcagtg tagagaaaaa tttcttatga ctcaaagtcc taaccacaaa 840
gaaaagattg ttaattagat tgcatgaata ttaagactta tttttaaaat taaaaaacca 900
ttaagaaaag tcaggccata gaatgacaga aaatatttgc aacaccccag taaagagaat 960
tgtaatatgc agattataaa aagaagtctt acaaatcagt aaaaaataaa actagacaaa 1020
aatttgaaca gatgaaagag aaactctaaa taatcattac acatgagaaa ctcaatctca 1080

gaaatcagag aactatcatt gcatatacac taaattagag aaatattaaa aggctaagta 1140

acatctgtgg caatattgat ggtatataac cttgatatga tgtgatgaga acagtacttt 1200

accccatggg cttcctcccc aaacccttac cccagtataa atcatgacaa atatacttta 1260

aaaaccatta ccctatatct aaccagtact cctcaaaact gtcaaggtca tcaaaaataa 1320

gaaaagtctg aggaactgtc aaaactaaga ggaacccaag gagacatgag aattatatgt 1380

aatgtggcat tctgaatgag atcccagaac agaaaaagaa cagtagctaa aaaactaatg 1440

aaatataaat aaagtttgaa ctttagtttt ttttaaaaaa gagtagcatt aacacggcaa 1500

agtcattttc atatttttct tgaacattaa gtacaagtct ataattaaaa atttttaaa 1560

tgtagtctgg aacattgcca gaaacagaag tacagcagct atctgtgctg tcgcctaact 1620

atccatagct gattggtcta aaatgagata catcaacgct cctccatgtt ttttgttttc 1680

tttttaaatg aaaaacttta ttttttaaga ggagtttcag gttcatagca aaattgagag 1740

gaaggtacat tcaagctgag gaagttttcc tctattccta gtttactgag agattgcatc 1800

atgaatgggt gttaaatttt gtcaaatgct ttttctgtgt ctatcaatat gaccatgtga 1860

ttttcttctt taacctgttg atgggacaaa ttacgttaat tgattttcaa acgttgaacc 1920

acccttacat atctggaata aattctactt ggttgtggtg tatatttttt gatacattct 1980

tggattcttt ttgctaatat tttgttgaaa atgtttgtat ctttgttcat gagagatatt 2040

ggtctgttgt tttcttttct tgtaatgtca ttttctagtt ccggtattaa ggtaatgctg 2100

gcctagttga atgatttagg aagtattccc tctgcttctg tcttctgaaa gagattgtag 2160

aaagttgata caatttttt ttctttaaat atcttgatag 2200

<210> 9
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> oligonucleodide primer

<400> 9
gcagtggctg aagatcaag 19

<210> 10
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> oligonucleotide primer

<400> 10
tgaggacgct gaccacac 18